DOCUMENT RESUME

ED 211 584 TM 820 005

AUTHOR	Bank, Adrianne; Williams, Richard C.
TITLE	Evaluation Design Project: Organizational Study. Annual Report 1980-1981.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	National Inst. of Education (ED), Washington, D.C.
PUB DATE	1 Dec 81
GRANT	NIE-G-80-0112
NOTE	237p.
EDRS PRICE	MF01/PC10 Plus Postage.
DESCRIPTORS	*Board of Education Policy; Case Studies; Decision Making; Educational Improvement; Elementary Secondary Education; *Instructional Development; Program Evaluation; *Research Design; *School Districts
IDENTIFIERS	*Evaluation Utilization

ABSTRACT

This report is arranged to display and explain the work of the Evaluation Design Project staff during the 1980-81 grant year. The intent of the entire project is to examine existing district policies, procedures and programs that appear to contribute to systematic use of evaluation and testing for educational improvement, and to make recommendations based on this examination for district guidance. Section I contains an explanation of the first two years of work and the intended efforts for the third. Section II presents a projected framework which emerged after the first year's study of four school districts and was refined after additional case studies in the second year. A draft of the paper "Considerations in Deciding on a District Management Strategy for Linking Testing and Evaluation Practices with Instructional Change" follows Secion II. The remaining five papers explore, from different perspectives, the categories presented in the framework section using data from the work field. The Appendix contains extended descriptions of two districts. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ter for the Study of Evaluation

UCLA Graduate School of Education
Los Angeles, California 90024

ED211584

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 820 005

DELIVERABLE - DECEMBER 1, 1981

EVALUATION DESIGN PROJECT:
ORGANIZATIONAL STUDY

ANNUAL REPORT 1980-1981

Adrianne Bank and Richard C. Williams
Project Directors

Grant Number
NIE-G-80-0112
P5

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California - Los Angeles

Acknowledgements

This document would not have been possible without the dedicated hard work of project staff including Michele Marcus, Pamela McGranahan, Donna Mitroff and Magdala Raupp. We would like to thank, also, others at CSE who discussed these ideas with us: Marvin Alkin, Eva Baker, Jim Burry, Don Dorr-Bremme, Joan Herman, Edys Quellmalz and Ken Sirotnik. Additionally, we wish to express our appreciation of Katharine Fry for the patience and good humor with which she tolerated seemingly endless revisions.

The six districts who made our work possible must remain anonymous; but knowledge of their efforts in the face of the turbulence surrounding public education will, we know, be helpful to many others.

The final responsibility for this document and for the work it represents remains with the project directors.

TABLE OF CONTENTS

	Page
Organization of Annual Report: 1980-1981	i
SECTION I	1
Explanation of the Evaluation Design Project: Organizational Study	2
Notes on Field Work Methodology	12
Work in Progress	20
SECTION II	32
Projected Framework for Analyzing a School District's Management Subsystem for Linking Testing or Evaluation with Instruction	33
Considerations in Deciding on a District Management Strategy for Linking Testing and Evaluation Practices with Instructional Change	38
Linking Testing and Evaluation Activities With Instruction: Can School Districts Make It Happen?	50
How Do External Factors Influence School District Management: A Preliminary Inquiry	65
School District Management Strategies to Link Testing with Instructional Change	84
Assessing the Effects of District Testing and Evaluation Efforts	101
Managing Teacher Behavior: From Symbolism to Reality	124
APPENDIX	146
Crescent City	1
Bordertown	37

Organization of Annual Report, 1980-1981

This document is the Annual Report of the second year's work of the Evaluation Design: Organizational Study project conducted at the Center for the Study of Evaluation.

The Report is arranged so as to display as well as explain the work of the project staff during the 1980-81 grant year. The reader of this Report should understand that many of the individual entries were prepared for specific purposes or publications prior to their being assembled in this Report. Therefore, there may be a jagged fit among the chapters if they are read consecutively. However, we have tried to smooth the transitions between and within sections as much as possible by some limited editing and the addition of introductory statements.

Some background about the project's past, current and anticipated work is essential to provide the reader with a context for reviewing the material. Section I contains an Explanation of the first two years of our work and our intended third year efforts. Also included is Notes on Field Work Methodology which describes some of the methodological issues which arose during the conduct of our second-year case studies. The third piece in this part is a description of Work in Progress.

Section II presents a Projected Framework which emerged after the first year's study of four school districts and has been refined after the second year's additional case studies. It is likely to be further clarified as the result of the Working Conference and the third year's practitioner-researcher collaboration on specific management problems connected with linking testing or evaluation with instructional change. This Framework is followed by the draft of a paper "Considerations in Deciding on a District Management Strategy for Linking Testing and Evaluation Practices with Instructional Change." The remaining 5 papers in this section explore from different perspectives the categories presented in the Framework using data from our field work.

The Appendix contains extended descriptions of two districts here named Crescent City and Bordertown.

Section I

Explanation of the Evaluation Design

Project: Organizational Study

This paper was originally written as a basis for discussion with the National Institute of Education about the third year work of the project. A version of the paper will appear in the winter volume of Evaluation News.

Explanation of the Evaluation Design Project Organizational Study

Adrianne Bank and Richard C. Williams

The Evaluation Design Project conducted by the Center for the Study of Evaluation, with Adrianne Bank and Richard C. Williams as co-directors, will have completed two of its three projected years of work by December 1, 1891.

The intent of the entire three-year effort is "to examine existing district policies, procedures and programs that appear to contribute to systematic use of evaluation and testing for educational improvement and make recommendations based on this examination for district guidance." (Three Year Plan, 8/30/79, p.6.)

Two years of field work in six "heroic" districts who have evolved management strategies to link testing or evaluation strategies with instructional change will be combined with practitioner contributions in a theory-into-practice Guidebook for managing data-based instructional renewal to the distributed to district and school personnel via commercial publication after the end of the third year.

Year # 1. During the first year of our work, we identified, through an extensive nomination procedure, four districts whose administrators, when initially interviewed, said the district was attempting to use test or evaluation data about student achievement as a guide to revising one or more aspects of its instructional activities. During subsequent interviews with district and school personnel in each of the districts, we found a variety of activities relevant to the linkage of testing or evaluation with instruction. For example, in one school district, staff development was intended as the key mechanism for school improvement. So, the patterns of childrens' scores on tests were examined to identify high priority content for staff development courses. In another district, far along in the process of coordinating a criterion-referenced testing system with a guided course of study and with teacher support materials, the learning specialists were examining the possible negative consequences of focusing student and teacher attention predominantly on basic skills. In yet another school district, we found that decentralized school site planning based on the results of standardized norm-referenced tests was being augmented (by court order) with a separate criterion-referenced testing program for racially isolated minority schools. Thus two separate district management sub-systems for linking testing and evaluation with instruction were emerging.

We found that not only did the specific activities connecting testing and evaluation with instruction differ from district to district, but so did the district's intraorganizational structures and arrangements. In some districts, each operation or service was focused on the testing/evaluation/instruction nexus. For example, budget and personnel directors in one district always examined their work in light of testing/evaluation/instructional needs. In another district, neither testing nor evaluation nor instruction were uppermost in people's thinking but instead were regarded as ancillary contributors to the district's thrust towards establishment of fundamental schools.

Finally, we found that the sequence in which the management system linking testing and evaluation with instruction evolved differed from district to district. But in no case did a district begin with a master plan to create such a management system. Rather, district personnel responded to immediate mandates from the state to develop tests, for example, or to the imminent availability of federal funds for some kind of project. Only afterwards did they rationalize what they were doing in such a way that their next efforts could be justified or understood in light of these recent experiences.

By the end of our first year's work, we had evolved a tentative set of common descriptors with which to view the phenomena we were interested in. We defined a T/E/I management sub-system to include those coordinated arrangements among district operations and services that affected data-based classroom instructional decisions. Some of our descriptors were: relevant environments, ideas, operations, coordinating mechanisms, impacts.

Year #2. During the second year, we investigated two additional school districts using revised interview instruments. We wanted to confirm and sharpen our understanding of factors and problems common in all school districts trying to link testing and evaluation with instruction and their unique manifestations in particular locales.

We identified two models by which districts linked testing or evaluation data with instruction. We characterized one as centralized, the other as decentralized. The centralized approach generally used criterion-referenced tests closely coordinated with a curricular scope and sequence. With the centralized approach operations and services such as staff development, media, budgets, and definitions of the roles and responsibilities of principals and other supervisory personnel, proved to be either tightly coupled or moving in that direction. This tight coupling seemed to be necessary for the effective implementation of a test-analyze test results-analyze instruction-revise instruction-reteach-retest cycle.

On the other hand, the decentralized approach used either norm-referenced tests or other evaluative information such as attitude surveys, or data on demographic changes. This data was given to individual schools by personnel within the evaluation unit. Such feedback was intended as part of the input which school site planning teams used as they reviewed the previous year and considered changes for the subsequent year. This arrangement—which left most of the initiative for instructional activities in the hands of the local site--seemed to conform with the loosely coupled administrative structure that we have recently come to associate with school organization.

Although one might expect that the size of the district would play a big part in determining the centralized versus decentralized approach, this did not appear to be the case. One small, as well as one large district had evolved a highly centralized CRT testing system with concomitant support systems. One small as well as one large district had moved toward the decentralized school site planning approach.

During the second year of our study, we also became interested in further analysis of the phases and the activities which characterized the evolution of the testing/evaluation/instructional management subsystem. In all of our districts, sequences of activities could be re-interpreted, through hindsight, as logical and intended; however, people in the district themselves characterized the process with terms such as confused, muddled, inefficient. They noted that when they were living through the process they did not have in mind the longer view or all the pieces of a larger conceptualization. They usually undertook their own activities in response to immediate problems, crises or requirements.

Someone said to us, "We were writing the book as we went along."

By the end of our two years of work, we have produced:

- six case studies of districts who are attempting to link testing or evaluation with instruction

- two annual reports. The 1980 report includes a literature review, an initial conceptual framework, and four case studies. The 1981 report contains a revised conceptual framework, a series of articles dealing with inter-district comparisons and two additional case studies.

- four papers. The first, "School District Linking of Evaluation and Testing with Instruction: Problems and Prospects," will be published by Sage in a book edited by Carol B. Aslanian called Improving Educational Evaluation Methods: Impact on Policy. Two others were delivered at AERA, 1980 and will be sent to professional journals within the month. They are "Linking Testing and Evaluation with Instruction: Can School Districts Make It Happen?" and "Testing and Evaluation as Strategies for Change." A fourth paper is being prepared for delivery at the Evaluation Research Society meeting in Austin; it is titled "School District Management Strategies to Link Testing With Instruction."

- a CSE monograph titled Evaluation in School Districts: Organizational Perspectives. This monograph contains four papers written by University of California sociologists who participated in the survey study which preceded research effort (Lyons, et al., 1978) and who are familiar with the six current case studies.

- a method for working collaboratively with representatives from participating districts.

- a Working Conference with district representatives and county representatives, professional association representatives, to discuss common issues and problems in managing testing and evaluation systems that impact instruction.

- a tentative outline for technical assistance materials to be developed by us and our school district partners.

Intended Third Year. In the past two years, we have come to realize that there are a variety of circumstances propelling school districts into establishing and maintaining testing or evaluation systems which impact on instruction.

These circumstances include: a movement toward competency-based testing in more than 35 states; an increasing interest in mastery learning; an enlarged capacity to carry out technically adequate criterion-referenced testing. It appears to be true, however, that attention to the management aspects of initiating, implementing and supporting a system which links data with decision-making--which connects testing or evaluation with instruction--has lagged behind attention to the technical aspects of test development or analysis and new evaluation metholodogies.

During our two years of work, we have discovered that there are few guidelines to assist districts in answering management questions such as the following:

1. START UP. How do you decide whether a data-based instructional improvement system is feasible in your district? Where do you begin?
2. COSTS. How do you estimate in advance how much it will cost in time, hiring, missed opportunities? How do you track the costs? Find the money?
3. MANAGEMENT. How do you utilize support and evaluation for teachers, principals, parents, board members? What other school and district operations and services need to be coordinated with such a system?
4. INSTALLATION AND IMPACT. How do you know whether operations are in place and working? How can you deal with problems, frustrations, grievances? How can you assess intended and unintended consequences; become aware of undesirable side effects?
5. SPECIAL GROUPS. Can data-based instructional improvement work at the secondary level? Is it useful for remediation, for average, above-average students? Appropriate to all?
6. AVAILABLE RESOURCES. What techniques and technologies are already in use. Should you adopt, adapt, or develop?
7. MAINTENANCE & RENEWAL. Is there a life cycle for data-based maintenance systems? How does renewal occur?

We have found that each district has idiosyncratically developed the implementation and maintenance aspects of whatever management strategy it is using for data-based instructional improvement. This responsive rather than pre-ordinate evaluation within districts seems not to be the result of inadvertence or lack of knowledge, but rather derives from the reactive nature of school district functioning in relation to such forces as community or agency pressure, availability of funding, initiative by committed leadership, etc.

Our third year activities have a single goal: to consolidate our own research--with reference to the relevant research from fields such as educational administration, educational change, dissemination--into a useful form for practitioners. We intend to create a theory-into-practice Guidebook which contains resource materials useful to any district where incentives already exist to develop testing/evaluation/instructional linkages.

We do not intend that our Guidebook provide a blueprint containing recommended steps-in sequence. Rather, it will address issues of importance, alert individuals to problems they might encounter, provide examples of how other districts have dealt with similar situations.

In order to develop a Guidebook at the appropriate level of specificity, in language familiar to district personnel, and containing useful information, we will engage in three types of partnership activities: A Working Conference at which an outline and format for the Guidebook will be discussed; CSE-district involvements with management problems related to data-based instructional improvement initially selected so as to be of interest to the specific district as well as to a larger number of districts; and contributions by practitioners of tips, techniques, anecdotes, and examples.

Prior to the beginning of the new contract year, we will have held the Working Conference for representatives of approximately twelve districts, county office, state department and professional association personnel interested in translating testing and evaluation data into instructional changes. At this conference our project will have presented our research findings, will have encouraged the exchange of information among district offices, will have identified special high interest topics within the overall subject of managing data-based instructional changes and will have developed a preliminary table of contents for the resource Guidebook.

Immediately after the start of the contract year, in December and January 1982, we will select three to five volunteer districts, perhaps from among those who attended the working conference. We will structure a collaborative arrangement with each district so that jointly we deal with a particular problem that the district has in managing the linkage between testing or evaluation and instruction. The three to five districts will be selected to represent salient problems encountered by many districts. In selecting the districts we want representation of a range of context factors: demographics, ethnicity, size, community characteristics.

At this time we expect that the Guidebook will contain two sections. The first will discuss generic issues relating to the management of data-based instructional change. It will deal with problems such as mobilizing and maintaining support for such ideas, using existing resources, coordinating, monitoring and reviewing linkage arrangements. A second section will contain illustrative materials from school districts' experience, tips and techniques around high priority items and perhaps anecdotes based on districts' experiences. We anticipate that parts of the second section will consist of contributions by central office administrators, principals, or teachers who have attempted to grapple with the issues.

From January to June 1982, we will be writing material; working collaboratively with districts on a consultative basis and reflecting on that collaborative experience; and soliciting reviews from other districts and researchers. In June through September 1982 we will be finalizing the content and text of the Guidebook. During this time, we will be producing several related articles for practitioner and research journals as well as a chapter dealing with our findings for an NIE-sponsored book on school district uses for evaluation. Our final report will be submitted in November 1982 for review, with revisions made by December 1, 1982.

The managers' Guidebook will have the following characteristics:

- it will be the first Guidebook of its kind addressing specifically the management needs as distinct from the technical aspects, of mounting a district wide effort linking data to instructional decision making. Such efforts are currently underway under various names: competency-based testing, mastery learning, school site planning, school improvement, etc.;
- it will bring together in a form useful to practitioners the results of these years of NIE funded project work, incorporating also the two previous years of NIE funded CSE efforts on school district evaluation units;
- it will be unique in that university and district staff will be partners in formulating, preparing and writing the Guidebook;

References

Bank, A. "School District Management Strategies to Link Testing with Instruction." Evaluation Research Society, Austin, Texas, October 1981.

Bank, A., and Williams, R. C. "Testing and Evaluation as Strategies for Change." AERA, Los Angeles, California, April, 1981.

Bank, A., and Williams, R. C. (eds.). Evaluation in school districts: Organizational perspectives. CSE Monograph #10. Los Angeles: University of California, 1981 (in press).

Lyons, C. D., Doscher, L., McGranahan, P., and Williams, R. C. Evaluation and school districts. Los Angeles: Center for the Study of Evaluation, University of California,]978.

Three Year Plan Proposed to the National Institute of Education. Los Angeles: Center for the Study of Evaluation, University of California, 8/30/79.

Williams, R. C., and Bank, A. "Linking Testing and Evaluation Activities with Instruction: Can School Districts Make It Happen?" AERA, Los Angeles, California, April 1981.

Williams, R. C. and Bank, A. "School district linking of evaluation and testing: Problems and Prospects." In Improving educational evaluation methods: Impact on policy. Carol B. Aslanian (ed.). Beverly Hills: Sage, 1981.

Notes on Field Work Methodology

This paper presents a discussion of issues which arose in the conduct of field work in this year's case study of large school districts. These issues are not unique to our study; many have been described in the literature dealing with qualitative research. Nonetheless, it seemed important to us to present them to illustrate both the strengths and the shortcomings of our research and analysis of the two districts described in the Appendix. We believe that field work in large organizations such as school districts or state departments of education will be enhanced by full discussions of the practical problems encountered.

Notes on Field Work Methodology

Adrianne Bank and Richard C. Williams*

During 1980-81, we were committed to additional field work in a number of school districts. Because of rising travel and personnel costs, the NIE project director agreed with CSE project directors that two additional districts would constitute sufficient new sites. The following is a brief analytic account, organized chronologically, of our field work activities.

Selection of Additional School Districts

Selection of the two additional school districts was to be based on several criteria: geographic location, travel accessibility, community and school district characteristics. An important consideration was the relationship of these two new school districts to the four school districts studied earlier in terms of the balance they are able to provide to the total set. For the purpose of providing a national emphasis, school districts outside of California were desirable. In terms of size and community environment, medium or large urban districts were desirable.

Based on nominations from CSE colleagues, other research colleagues, and selected school and government officials, twenty-six school districts were identified as possible additional research sites. Initial school district screening was conducted through telephone interview with each district. The screening questionnaire included questions regarding the community (urban, suburban, or rural), school level demographics (the size of the school district, ethnic make-up, achievement levels), changes in the community and/or district that might affect testing and evaluation policies, and the districts' efforts relating T/E/I. The school districts' willingness to participate in the study was also ascertained. The field of possible school sites was narrowed to ten districts. Further screening interviews brought the final selection to two districts.

---

*The authors thank Emily Brizendine for her assistance.

Bordertown and Crescent City met all our requirements. CSE had previously studied Bordertown; so we had background from which to look at current evaluation and testing activities. We selected Crescent City because it is unique in some respects: The district is large, the communities it serves are both rural and urban, and it has established a tightly-coupled system of linking evaluation and testing to the instructional component. It was also characteristic of other school districts because of its size and its problems in coping with population shifts and high pupil transiency.

Issues. We realized that our total sample of only six school districts required us to be most careful about making generalizations from our findings. It was clear from our first year's work that each district was a unique organization of people and functions influenced very heavily by local environmental pressures. We wanted the two new districts to add to the heterogeneity of the mix. At the same time, we wanted to investigate how they grappled with the turmoil common to most districts at this time.

Entry to School Districts

We sent a letter to each district's Superintendent notifying him of the district's selection. In addition, we phoned the head of the research and development units of the districts. In further correspondence, we outlined what we proposed to accomplish and outlined ways in which we would like to work with the districts. In one case, the approval process for our study was quite formal. We drafted a document explaining our purpose and procedures. In the other case, phone conversation was sufficient.

Bordertown was interested in participating because we had worked with them previously; they felt it was important because of NIE's interest in the study, and they desired feedback on their district's program. Crescent City's willingness to participate came from their shared interest in our research concern of linking testing and evaluation to instruction. In addition, they desired feedback from us on their own system of T/E/I.

Issues. In both districts, we received the tacit support of the Superintendent and the active involvement of the person responsible for evaluation and testing. In neither district did our study arouse great anxiety as to what we might find or to whom we would report our results. This was in marked contrast to concern about confidentiality raised by the smaller districts in our first group. It may be that in larger districts, reports or studies are so frequent, and so often are of very little consequence that, realistically, the evaluation people did not have much concern. Neither did they ask to see copies of the interview questions or of the final report. Although they were interested in the topic of inquiry, they were surprisingly passive in asking us how what we were learning might be of benefit to them.

Identifying Respondents

A letter was sent to the two school districts indicating our interest in interviewing central office personnel, those in peripheral functions related to testing and evaluation including board members, and teachers and principals. The school districts identified the respondents for us and set up interview schedules. For Crescent City, our contact person was the Director of Research and Development, who selected individuals for us to interview during the four-day interview period. We determined which members of our three-person research team would interview whom. In contrast, our contact person in Bordertown, the head of the evaluation unit, selected the respondents and arranged the interview schedule in such a way that each member of the interview team was designated to interview certain people at certain levels of the school organizations.

These somewhat different approaches had an impact on our field work. For example, in Crescent City we were able to have each research team member pursue one or more lines of inquiry, e.g., R&D, curriculum, testing. In Bordertown, however, each researcher talked with respondents ideosyncratically across topics. In the second case study we had less of a comprehensive view of major components of the district's activities.

Issues. In large school districts, researchers have problems in obtaining and in knowing that they've obtained representative respondents. We have no reason to suppose that our interviewees were stacked to represent a particular point of view. Rather, the schedules were made up on a catch-as-catch-can basis. It was particularly difficult to get teachers and principals free of their site responsibilities to talk with us. We did not have, in advance, background data on our respondents. We therefore knew little about the representativeness of their opinions or the perspective from which they spoke Some of that we gleaned from the interview itself. But, our interviewing team felt dissatisfied with this middle level of our involvement in the district. We were spending a good deal of time getting to know the districts and speaking with over forty people in each; nonetheless, we were not able to satisfy ourselves that we had tapped into the most important factors operating in that district.

Interview Questionnaires

There was a great deal of discussion among project staff as to how to best manage the information yielded from the interviews. In order to get consistent data, the interview questions were based upon those from the first year study. In formulating the questionnaires, our primary concern was how to obtain an accurate picture of the situation at these two large school districts within the limited amount of time available per interviewee. We decided to modulize the questions according to subject areas and key them to the analytical framework.

Issues. Within the project staff there were two divergent orientations. Some of us wanted a structured set of questions and a given order for the questions so as to get comparable data across respondents. Others of us believed that the order of introducing the topics could influence the "story" we were getting and that we would get a broader picture by letting individuals tell us what they knew in their own way. In view of our limited time with respondents, we reached a compromise. We sorted our questions: some we would ask of everyone, others would only be asked of people who we believed could have important things to tell us in that area.

Preparing Team for Case Study

Field research training sessions were held for all project staff. At those sessions, question ambiguities were clarified, the study's intent was explicated and note taking and interviewing techniques were practiced.

Issues. We were fortunate that all interviewers were also researchers of the project staff. We only needed to informally review, rather than train, ourselves for the task. On the other hand, personal note taking styles predominated in spite of before-interview agreement on how to record data. Only time-consuming and highly prescriptive training sessions--which we did not believe worthwhile--would have assured consistency in the amount of specificity, examples, length, etc. of notes.

Collecting Data and Analyzing Results

While the length of the interviews varied, the interviews averaged about 45 minutes. In spite of our agreed-upon procedure of systematically asking questions in a certain order, the interview team members found themselves, in practice, having to modify the procedures as time and the situation dictated. For example, in both districts, the Superintendent of Schools was an important respondent. However, he was unable to meet with us for more than half an hour. Thus, we had to rather quickly adjust our interview to address with him those issues about which he would have special knowledge or insights.

Issues. During the field site visits, a number of methodological problems emerged. One problem resulted from the impact that time limitations, the schedule format and the intensity of the interview period had on the interview team. In the process of interviewing district personnel, it became difficult for the interview team members to coordinate and integrate the data being accumulated and make the necessary adjustments to the procedure as it progressed. Because of everyone's familiarity with project interviews, individual interviewers were able to make appropriate modifications in the pre-arranged plans. However, we suspect that this problem is common with interviewing teams. Our feeling was that we might have scheduled fewer interviews with more analysis time; on the other hand, our need to take advantage of being on-site and having access to people made this choice difficult.

Another problem was the recording of the interviews. We both took notes and made a tape recording for each interview. Sometimes mechanical problems (i.e., malfunctioning batteries) caused the amount of data obtained to differ from interview to interview.

Report Writing

The writing of each final report was designated to two members of the project staff. They examined our interviews, organized the report, shared the drafts with others who were members of the same interview team for comments and then re-wrote the report. Before the case reports are finalized, they will be reviewed by district personnel.

Issues. Each of the case studies is approximately thirty pages in length. It would have been possible to double or triple the number of pages to provide readers with quotes and detail. Such a level of specificity would have enabled readers to form their own hypotheses or conclusions about the districts' linkage strategies. On the other hand, much shorter versions in which we selected only that information clearly relevant to our interest would have made our analysis crisper. We chose what we hope is a middle-ground--enough for the non-specialist interested reader to understand what is going on in the district with regard to linking testing and evaluation with instruction.

Work in Progress

Summarized here are papers and other work undertaken by the project during the 1980-81 year which will be completed during the 1981-82 year. They include:

1) CSE Monograph # 10: Evaluation in School Districts: Organizational Perspectives. Adrianne Bank and Richard C. Williams, (eds.).

2) Evaluation as a Decision-making Tool For School District Administrators. Richard C. Williams and Pam McGranahan.

3) Evaluation Comment Special Issue: Linking Testing and Evaluation with School District Instructional Programs. Proceedings of CSE Working Conference, November 5-6, 1981.

CSE Monograph #10: Evaluation in School Districts: Organizational Perspectives. Adrianne Bank and Richard C. Williams (eds.)

This monograph consists of four papers by University of California sociologists who used the CSE 1979 national survey of school district evaluation units to reflect on the relationship between evaluation activities and the districts' organizational contexts.

This monograph has grown out of CSE's on-going interest in the conduct of educational evaluation, how it operates on school districts, and how it contributes to educational practice. The monograph's title--Evaluation in School Districts: Organizational Perspectives--we believe, conveys the importance of its thematic context. Three elements drawn from this title serve both to introduce the volume and to underscore its principal themes.

First, the monograph focuses on evaluation and how it occurs within a specific setting--that of the central administrative offices of school districts. This narrows the interest from evaluation-in-general to evaluation-in-operation within this special context. Second, the volume examines the relationships between evaluation and school district organizational features that impinge upon evaluation. Research on evaluation methodology and evaluation use often pays lip service to the need to understand the context within which evaluation occurs. But there are very few research studies which attempt to relate organizational variables such as decision-making, internal administrative structures, role definitions, and the nature of districts' "technical core" to the way in which educational evaluations are conceived, conducted, and perceived by district personnel. Thirdly, the monograph offers multiple perspectives on organizations rather than elaborating a single perspective.

Our understanding of organizations and of administrative decision-making has undergone important changes in the last 30 years or so.

In this monograph the authors refer to a wide range of new notions and concepts about organizations. Three such ideas occur quite often and the reader not acquainted with the organizational literature may not be familiar with them. The three concepts are "loose coupling" (Weick, 1976), "institutional" vs. "technical" organizations (Meyer & Rowan, 1977), and "organized anarchies" (March & Olson, 1976).

Loose coupling refers to the degree to which various parts of an organization (its individual staff, divisions, units) are linked to one another. If one organizational unit acts, to establish a new procedure, for instance, will it have the desired impact or effect on other parts? If not, the organization is loosely coupled.

Institutional and technical organizations differ on two dimensions. The first difference is that, while technical organizations have a well-defined technical core that can be rather clearly understood and to which the organizational structure is "tightly coupled," institutional organizations, such as school districts, do not.

The second difference between institutional and technical organizations is the degree to which they are held accountable to societal rules. The more an organization is supported by public monies, the more it is "institutionalized" and hence is judged by the expectations accompanying those monies. The evaluation unit of this kind of organization, such as a school, again, will devote its resources to demonstrating compliance to external expectations rather than with internal information needs.

"Organized anarchies" is a term used to describe the phenomena that actually occur in many organizations and which are quite counter to what rational bureaucratic theory would lead one to expect to find. This kind of behavior can be described as: "choices looking for problems," "solutions looking for issues," and "decision-makers looking for work."

If organizations, specifically school districts, have the characteristics embodied in the three concepts, then there are profound implications for thinking about and perhaps re-considering the role and value of educational evaluations in organizational decision-making. At the very least, it probably means that evaluators have to come to understand the organizational constraints within which administrators make decisions. This understanding is crucial if evaluation information is intended to be used in the decision-making process.

During the past fifteen years the education evaluation community has exploded into a profession. The evidence of its growth is everywhere.

As evaluation methodology is being developed by some, others begin to explore whether the resultant evaluation studies are contributing to decision making. This research suggests that evaluation does not seem to be having the desired effects. As a consequence, some evaluation specialists are redoubling their efforts to further refine evaluation techniques. Others seek to proselytize about the value of evaluation. Still others take a closer look at the phenomenon of evaluation utilization or non-utilization.

Those who are interested in evaluation utilization can be grouped into several loosely-linked categories. Although they are all interested in the same general idea, their individual perspectives give their work different emphases.

There are those scholars who see the study of evaluation utilization as a special case of the study of knowledge utilization (Caplan, 1979; Weiss, 1979; Baruch, 1980). They are concerned with how the production of research knowledge in specialized fields, including education, gets transmitted in a useable form to those who must make policy decisions. The audiences of principal concern to these researchers are primarily, though not exclusively, federal and state legislators and administrators who must allocate scarce resources among competing program alternatives.

There are also individuals working on evaluation utilization from the point of view of defining the problems inherent in the practice of evaluation and improving its procedures and practices (Alkin, 1979; Patton, 1978). These researchers view evaluation as a service to decision-makers, and so they reason that if such services are perceived as useful to clients--at whatever level the clients may be--then the evaluation profession needs to reexamine the way in which it goes about its work.

A third group (Appling & Kennedy, 1980; King, Thompson & Peckman, 1981; Williams & Bank, 1981) approaches the evaluation process and the data it generates as one of many sources of information residing within a dynamic and often politically volatile context. They examine evaluation as a process and they consider the use or non-use of evaluation findings in local districts in terms of the logic of these organizational settings.

It is only within the past five years that this last group, who are interested in examining the organizational (as distinct from the knowledge production or the professional) aspects of evaluation utilization, have become a major force within the evaluation research community. We believe that CSE work, underwritten by NIE and expressed in this monograph, has made an important contribution to this development.

References

Alkin, M. C., Daillak, R., White, P. Using evaluations: Does evaluation make a difference? Beverly Hills, California: Sage, 1979.

Apling, R., Kennedy, M., & Newman, W. F. The role of evaluation and test information in public schools. Cambridge, Mass.: The Huron Institute, 1980.

Bank, A., & Williams, R. C. (eds.). Evaluation in school districts: Organizational perspectives (CSE Monograph #10). Los Angeles: University of California, 1981 (in press).

Boruch, R. F., & Cordray, D. S. (eds.). An appraisal of educational evaluations: Federal, state and local agencies. Northwestern University, 1980.

Caplan, N., et al. The use of social science knowledge in policy decisions at the informal level. Ann Arbor: Institute for Social Research, University of Michigan, 1975.

March, J. G., & Olsen, J. Ambiguity and choice in organizations. Bergen, Norway: Universitetsforlaget, 1976.

Meyer, J. & Rowan, B. Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, 1977, 83, 304-363.

Patton, M. Q. Utilization-focused evaluation. Beverly Hills, California: Sage, 1978.

Thompson, Bruce, & King, J. A. "Evaluation Utilization: A Literature Review and Research Agenda." Presented at AERA, Los Angeles, 1981.

Weick, K. E. Educational organizations as loosely coupled systems. Administrative Science Quarterly, 1976, 21, 1-9.

Weiss, J. A., & Weiss, C. H. "What Makes Social Science Useful for Public Policy?" Presented at the annual convention of the American Psychological Association, New York, September, 1979.

Evaluation as a Decision-making Tool For School District Administrators

Abstract*

by

Richard C. Williams
Pam McGranahan

The evaluation community has invested considerable time and energy in investigating whether or not decision makers have used evaluation research. Based on assumptions and empirical evidence that evaluation utilization has been minimal, practicing evaluators and evaluation academics have derived several solutions that would, from their perspective, increase utilization. Among the solutions are:

- Better staff. In a CSE national survey of evaluation unit directors a large percentage of respondents commented that they need more and better trained staff if they are to turn out high quality work which would meet decision makers' demands.
- Better training. Evaluations are only as good as the quality of the evaluators who conduct them. The field is growing and in need of high quality, agreed-upon methods and approaches. Accordingly, a large-scale cooperative effort of evaluators produced Standards for Evaluations of Educational Programs, Projects, and Materials.
- Better understanding. Under-utilization, it is agreed, results primarily because evaluation reports are poorly timed or do not address salient administrative concerns. This can be corrected by having evaluators become more sensitive and responsive to decision makers' needs.

While there is, likely, merit to these concerns and recommendations, we doubt that they fully address the problem of school district under-

---

* When completed, this paper will be submitted to a journal read by administrators or by school board members.

utilization of evaluation reports. All of these recommendations are based on some fundamental assumptions which the evaluation community accepts but which need, in our opinion, some careful attention. These assumptions are:

1. Administrators generally value data-based evaluation research.
2. Evaluation can provide utilizable data-based information.
3. Data-based evaluations are instrumental in decision-making.

While the above assumptions are likely true in some instances, there are numerous occasions when they are not. And increased resources or better trained evaluators may not result in increased evaluation utilization.

For example, there may well be times when a data-based evaluation report is not needed; indeed, the administrator may consider such a report dysfunctional. The administrator may be in a tenuous political situation, or he or she may have legitimate reasons to continue a project that have little relationship to whether or not the project has achieved certain goals.

Evaluations of many programs are severely limited in terms of the indisputability of their findings. There are several reasons for this: evaluations are generally conducted in the field under non-experimental conditions. While such studies can be useful they can be easily discredited by political foes. Also, especially in instructional evaluations, the technical core of what is being evaluated is often weak and this makes a sound, defensible evaluation difficult to achieve.

Finally, decision makers have many bases upon which they make program decisions, such as: personal preferences, peer group pressure, advice from trusted colleagues, misunderstanding of data or information provided. Data-based evaluations may play a pivotal role in decisions; or they may have little or no impact depending on the configurations of circumstances surrounding the decision.

We suspect that the future direction of evaluation units will be profoundly affected in the next few years. If federal block grants, rather than categorical grants, become a fiscal reality, school districts will likely have greater discretion in deciding on whether or not limited resources should be spent on continuing evaluation activities. We hope school districts will think deeply about the potential uses such units might have. If they merely extrapolate from previous district evaluation units activities, many districts may decide that evaluation units are not worth the money. Such a decision may be unfair to units that may not have been very useful to local needs because such units have had to direct their attention to satisfying mandated external funding agencies.

With the presumed decreased need for external reporting and monitoring it may well be a rare opportunity for school decision makers to develop a decision-relevant evaluation unit. We suspect, however, that this will not happen unless decision makers make a concerted effort to shape school district evaluation programs and units so that they serve specific school district needs. Administrators must assure that the evaluation unit is addressing the correct problems and reporting the data, whether favorable or not, in a way that is useful. We sense that evaluators, concerned as they are about lack of impact, will be receptive to such administrative overtures. Out of such mutual needs may arise a new generation of school district evaluation units that indeed service local school decision makers' wish for relevant information.

Evaluation Comment

"Linking Testing and Evaluation with School District Instructional Programs"

Proceedings of the CSE Working Conference

This special issue of Evaluation Comment, mailed to the regular 10,000 subscribers as well as several hundred district personnel indicating interest in the topic, will emphasize the productive collaboration occurring between CSE researchers and district administrators.

The Comment will include abstracts or summaries of the following presentations:

Adrianne Bank and Dick Williams, CSE Project Directors, on "Themes and Variations in Six School Districts";

Marv Alkin, CSE Project Director, on "Evaluation Use in Local Schools: Progress Report on CSE Studies of Schools";

Joe Felix, Director of Evaluation, Cincinnati, on "Local School Budgeting: A Focus for Evaluation";

Joe Gastright, Assistant Director of Evaluation, Cincinnati, on "Expanding the Framework for Interpreting Norm-referenced Test Results in the Classroom";

Mary Kennedy, Project Director, Huron Institute, on "Things That Can Go Awry When Tests Are Used to Manage Instruction";

Jim Popham, UCLA Professor, on "Detroit's Minimum Competency Testing Program: A Catalyst for Instructional Improvement";

Bill Spady, Director, National Center for the Improvement of Learning, on "The Evaluation and Credentialling of Children: The Tail That Wags the Dog";

Theron Swainston, Associate Superintendent, Las Vegas, on "Linking Testing and Evaluation to a Total School Management System From the Classroom to the School Board."

It will also contain a digest of the reports of more than a dozen school districts about their activities in testing and evaluation followed by an analysis of these activities using the analytic framework developed by the Evaluation Design: Organizational Study Project. Additionally, the "Promising Practices" compiled at the Conference by small groups interested variously in start-up, costs, resources, maintenance and renewal of management subsystems, will be listed. A report will be made on the Conference discussion concerning the content and format of the Management Guidebook to be developed with practitioners by CSE during 1981-82.

Section II

Projected Framework for Analyzing a School District's Management Subsystem for Linking Testing or Evaluation with Instruction*

This framework now consists of a list of elements important in describing existing or prospective management subsystems which link testing or evaluation with instruction in school districts. To complete this framework requires the development of indices for each element and a specification of the relationships among the elements as they develop over time based on data from school district practice.

---

*This document is being distributed as CSE Technical Report #156.

Projected Framework for Analyzing a School District's Management Subsystem for Linking Testing or Evaluation with Instruction

Adrianne Bank and Richard C. Williams

Description of the Proposed Framework

The framework to be developed will provide a basis for analyzing a school district's management subsystem that connects testing and evaluation activities with instructional change. The management subsystem that we are interested in is a construct--that is, it is not a single organizational entity such as a division or a unit that is clearly identified within a district, but such a subsystem, when it exists, does have properties which can be described and analyzed across districts. It is not necessary to an understanding of the projected framework to define precisely what we mean by the terms "evaluation," "testing," or "instruction." Common usage suffices for the present--except to say that we exclude from our definition activities which are initiated, or used by a single classroom teacher. We are also defining instruction broadly: to include curriculum, materials, teaching strategies, classroom management, etc. Instructional change may mean changes in policy, in management, in individual's or group's knowledge, skills or attitudes.

The framework now consists of a set of elements which may be important to understanding any management subsystem but which we have applied to the management subsystem linking testing or evaluation with instruction. We have identified and described these elements. Our next step will be to hypothesize relationships among the elements. We will also attempt to make profiles of specific testing/evaluation/instruction management subsystems using these elements and relationships. Finally, we will try to describe what we have observed about the dynamic interaction among these elements over time.

Development of the Framework

After collecting interview data and examining district documents we wrote a first draft case study for each of six districts which were engaged in activities related to linking testing or evaluation data with instructional change.

We then attempted to identify categories or elements of relevance across districts to see what analytic insights this might yield. Although the details of our districts' activities were unique, we wanted to infer functional equivalents across districts.

We looked at existing analytical frameworks, recategorized data by their headings and tried to define the relationships between and among the categories or elements.

We discovered that existing conceptualizations did not seem to fit what we had found. For example, we considered the standard planning or development model with its sequential processes--define the problem, search for solutions, select the best solution, implement the solution, evaluate the results, and then recycle the process. We did not find this process occurring in the districts even when we defined as the "problem" the linkage between testing, evaluation and instruction. Many researchers of credence supported our view that such a rational, sequenced decision-making cycle seldom, if ever, is actually utilized in planning and implmenting programs (Lindblom, 1959; March & Olsen, 1972; Clark, 1980).

We considered a related construct that applied to the stages of educational change, i.e., mobilization, implementation, institutionalization (Berman & McLaughlin, 1975). It too, did not apply even when we defined the management subsystem we were looking at as an "innovation." Generally, in our districts a management subsystem for linking tests or evaluation with instruction was not recognized as a change or innovation until after it was almost fully developed, if at all.

We considered various methods of force field analysis. The management of testing or evaluation for purposes of instructional change could not be conceptualized as a clearly defined or planned "program" that was buffeted by forces acting either for or against its use. Nor could we identify points in time during which the system went through "unfreezing" and "refreezing" in relation to constructs we were interested in.

The most useful general approach for us seemed to be open systems analysis. Open systems analysis is not a theory exploring specific sequences of causes and effects. Rather, it "is an approach and a conceptual language for understanding and describing many kinds and levels of phenomena" (Katz, D., & Kahn, R. p. 452). The approach stresses the following points (Hanson, E., p. 181-182):

- focus is on the way an organization does function rather than on the way it should function;
- there are many ways to perform a task that are equally satisfactory;

- an organization functions with dynamic, not static, relationships;
- the demands and needs of the environment give direction to organizational events;
- organizational leaders are subject to events that are not of their making and are beyond their control;
- organizational power is distributed into subsystems which must differentiate and integrate their activities;
- communication is designed to integrate the activities of subsystems and establish linkages with the environment.

School districts are complex organizations; thorough understanding of these organizations so as to locate a management subsystem within them in a full open-systems sense was beyond the scope of our work. We, therefore, began to think about elements of importance that would help us describe what we had observed.

Anticipated Work on the Framework

As we have already indicated, we do not yet have a completed framework. We only have a set of categories or elements which help us think, systematically, about the management problems of connecting evaluation and testing data to instructional chance. Some of the ways in which we hope to drive the framework forward include:

- suggesting indices for these elements. That is, listing the phenomena which could be observed to indicate the presence, absence and characteristics of the elements;
- suggesting relationships among elements in the framework and noting that these relationships change in direction and in influence over time;
- describing patterns or profiles which seem to occur in school districts.

Anticipated Uses for the Framework

- The elements can provide a common language in which to describe existing or projected subsystems. They can be used as a checklist to reveal where important considerations have been overlooked.
- The hypothesized relationships can be used as guidelines to assess the probabilities of success or of failure of possible courses of action related to the creation or maintenance of management subsystems.
- The profiles can succinctly describe actual rather than ideal practice.

Elements in a Framework for Analyzing Management Subsystems

Elements	Descriptors
Relevant Environments:	
types	societal, community, professional, intra-district, etc.
characteristics	stimulative, supporting, neutral, mixed, hostile
influence	strong, moderate, weak
Ideas:	
types	technical ideas; management ideas, etc.
characteristics	comprehensive/fragmented; consistent/inconsistent; diffused/isolated; consensual/controversial
idea-holders	pro-active/reactive; champions/supporters/neutral/negative
Operations:	
types	testing, evaluation, curricula, instruction, supervision, staff development, budget, etc.
degree of integration	central, peripheral, uninvolved
staff competency	high, medium, low
Coordinating Mechanisms:	
types	formal, informal; within/between levels
functions	motivate, communicate, decide, act, monitor
Impact:	
types	intended/unintended; direct/indirect; test-related/person-related
affected levels	community, board, central office school, classroom
affected parties	students, parents, teachers, principals, administrators, policy makers

References

Berman, P. & McLaughlin, M.W. Federal programs supporting educational change, Vol. IV: The findings in review (R-1589/4-HEW). Santa Monica, California: The Rand Corporation, 1975.

Clark, D. L. In consideration of goal-free planning: The failure of traditional planning systems in education. In D. L. Clark, S. Mckinnin, & M. Malkas (eds.), New perspectives on planning in educational organizations. San Francisco: Far West Laboratory, 1980.

Hanson, E. M. Educational administration and organizational behavior. Boston: Allyn and Bacon, Inc., 1979, p. 181-182.

Lindblom, C. E. The science of 'muddling through.' Public Administration Review, 1959, 19, 79-88.

Katz, D., & Kahn, R. The social psychology of organizations. New York: John Wiley & Sons, 1966, p. 452.

March, J. G., & Olsen, J. Ambiguity and choice in organizations. Bergen, Norway: Universitetsforlaget, 1976.

Considerations in Deciding on a District Management Strategy for Linking Testing and Evaluation Practices with Instructional Change

This paper represents a first attempt at translating the elements of the projected Framework into a specific set of considerations. Questions such as those in the paper can act as an early warning device to sensitize district personnel to present or potential problems in managing the connection between testing or evaluation and instruction. Checklists such as this will be developed in the third year of the project for inclusion in the intended managers' Guidebook.

Considerations in Deciding on a District Management Strategy for Linking Testing and Evaluation Practices with Instructional Change

Magdala Raupp

The intent of the Evaluation Design Project conducted at the Center for the Study of Evaluation, with Adrianne Bank and Richard C. Williams as co-directors, was to examine a small sub-set of existing district policies, procedures and programs that appear to contribute to systematic use of evaluation and testing for educational improvement. It was expected that the project by addressing itself to the influences on and the concerns of school and district personnel, would be in a position to make suggestions for other district's guidance.

Through an extensive nomination procedure, six districts were identified whose administrators claimed their district was attempting to use test scores or evaluation data as a guide to revising one or more aspects of their instructional activities. During subsequent interviews with district and school personnel in each of the districts, we found a variety of practices that districts used to link testing and evaluation with instruction. We found that the specific practices differed from district to district as did the district's intraorganizational structures and the sequence in which their management subsystem linking testing and evaluation evolved.

In the districts we visited, we saw practices that can be roughly grouped into three configurations. These configurations will be referred to as strategies and are fully described in the paper "School District

Strategies to Link Testing with Instructional Change", (Bank, 1981). We do not mean that these three are the only possible strategies nor that the districts set out to create a system to implement a particular strategy.

1. An instructionally-oriented, objectives-based strategy. This district adopted a structured diagnostic/prescriptive teaching supported by a district-wide scope and sequence outline of objectives, a criterion-referenced testing system continuously updated, materials cross-referenced to the objectives and to the tests.
2. A personnel-oriented staff development strategy. Great school-to-school variability and the likelihood that principal and teaching staffs would remain stable influenced at least one of the districts to adopt a personnel-oriented staff development strategy as a key to data based instructional change. Their assumption was that teachers themselves made the major difference in student learning and that data about deficits in student achievement should determine the content of staff development courses.
3. A building-oriented problem solving strategy. Schools in this district for reasons of ethnicity, geography and tradition represented distinctive organizational entities. The district felt that school staffs and parents should together identify their problems and devise solutions using testing and evaluation data.

Although the term strategy will bring to mind direction and purpose we found that most districts we visited had not made a plan or a blueprint prior to taking the actions that seemed to result in a strategy. What they did evolved out of events, interests of people, effects of the environment. At some point in time, these activities were conceptualized or reconceptualized so that future activities could be justified or made plausible. At that point, and for some districts, what could be called a strategy emerged.

A wide range of reasons influenced our six districts to do what they did to use data to influence instructional decisions. Clearly the immediacy of state or federal mandates and funds was one factor. Obvious shortcomings in student skills in some cases, internal pressures within particular district offices or the special interests of individuals in positions of power on the Board or within the district provided the impetus. Districts seemed to be, in this manner, adjusting and accommodating - in the Piagetean sense of the phrase - to the various inputs and demands made upon them.

Although what we described above was the pattern we most often found, we suggest that it is possible for leaders in districts to visualize and shape a management strategy in a way that is more proactive for subsequent activities. It should be possible for leaders to understand many of the constraints and influences active on and within the district and then take reasonable steps to move in an instructional change direction. If we believe public schools should be seeking instructional change thereby improving student learning, and if we believe testing and evaluation can contribute to this change, we should clarify, as best we can, the process of developing a testing/evaluation/instruction management subsystem that will be uniquely suited to the specific situation of a given district.

In this paper we will deal with some of the considerations that might go into a decision to use a strategy for instructional change in which data from testing and evaluation would play a major role. We will use, as organizers, some of the elements alluded to in the analytic framework (Bank and Williams, p.33). There are certain things these elements can do and other things that they cannot do. They can help us think about the district as an organization embedded in and responsive to its environment. They can bring attention to ideas, operations and mechanisms already in place. An analysis using these elements can make the decision making process somewhat less uncertain. Such analysis will raise but not answer our questions. The decision to develop a management strategy and the specifics of that strategy must be unique to a district and not acquired as a shelf-item from other districts. While it is true that many districts share common characteristics, in no two districts is the combination and arrangement of characteristics the same.

Commitment to develop a strategy for managing data based instructional change will lead district personnel into a multitude of decisions and considerations of complex issues. Since effort beyond that required simply to maintain the status quo is needed, a first major question might be: IS SUCH A T/E/I MANAGEMENT STRATEGY WORTH THE EFFORT? The answer to that general question may be forthcoming from answers to more specific questions.

- are administrators, teachers, parents satisfied with what the districts' schools are doing in relation to student learning?
- do these individuals see a gap between what the schools are doing and what it would be possible to do?
- do these individuals believe that educational leadership requires activism and a constant search for better alternatives?
- do these individuals believe that more effective instruction can come, at least in part, as a result of examining the present performance of students on tests or the present performance of programs as indicated by evaluations?

A second question, suggested by the element, relevant environments, might be: WHAT ARE THE OPPORTUNITIES AND THE CONSTRAINTS POSED BY THE ENVIRONMENT IN WHICH THE DISTRICT ORGANIZATION IS EMBEDDED? HOW WILL THESE ENVIRONMENTS SHAPE THE MANAGEMENT STRATEGY?

Relevant environments include factors external to the district organization but which act upon it. Geographic factors such as the existence of schools isolated or clustered together; community factors such as size, its socio-economic, political or religions organization; historic factors such as its past tradition or reputation and its expectations for the future. Such factors in the environment can help or hinder data-based instructional change. Answers to the specific questions below might help to answer the more general foregoing question.

- how have testing and evaluation results been historically used in this district?

- do parents and members of the district community usually share the belief that test results reflect the learning of students and that the district has responsibility for increasing that learning through instruction?
- what specific instances are there of this belief? Is it shared by all parents?
- what are the external pressures and incentives to use testing and evaluation results to improve instruction? have there been federal or court mandates? how strong a pressure has the media, through its critical appraisal of the publis school system, put on the district?
- what is the attitude of the opinion leaders in the community towards the issue of more efficiency in the schools and how supportive would they be in the event that the district commits itself to a strategy for improvement?
- geographically, what is the organization of the schools within the district? are the same issues being considered at the different schools or are the issues and problems different at each school site?

A third general question might be related to ideas, an element in the framework that refers to ideas specific to testing, evaluation, instruction and how they interact with one another; also included would be those ideas specific to the management of the subsystem linking testing and evaluation with instruction: WHAT IDEAS RELATED TO TESTING AND

EVALUATION RUN CONSISTENTLY THROUGH THE DISTRICT AND HOW WOULD THEY AFFECT CHOICE AND IMPLEMENTATION OF A STRATEGY FOR INSTRUCTIONAL IMPROVEMENT? Are there ideas that converge to a common vision valued by members of the staff, members of the community and parents? Group members have images of themselves and agree on the way in which they perceive the world, based on common experience, values and beliefs. Future action is going to be based on those ideas that members of groups share. It is important, therefore, to probe into the content and direction of these ideas.

- what ideas do the teaching staff have about testing and evaluation that would support or undermine the effort to use data generated by both for instructional improvement?
- what proportion of the staff would experience anxiety or threat regarding the use of test results to diagnose deficits in the instructional process?
- what proportion of the staff would be willing to undergo training in new skills and behaviors necessary to establish the test, evaluation, instruction link?
- what proportion of individuals - members of the community, parents, media people - would be willing to participate in the effort?
- is there a 'critical mass' or total number of individuals within the staff, members of the community and parents that are genuinely concerned over the issue of instructional improvement and would be supportive of measures taken in that direction?
- does this 'critical mass' see the proposed action as appropriate and adequate, that it will move the schools in the direction of the target, and that the benefits will outweigh the costs?

to which extent are those in leadership position able to create a common vision that is valued by many members and a sense of urgency, of needing to take initiative and action?

- are there 'idea champions' that could imbue the staff with improvement fervor? what is the personal style of those individuals, what makes them important and how could they become an asset to the strategy under consideration?

Another major question could be posed in relation to the operations element. This refers to those organizational units that are to be included in the management subsystem, such as, testing, evaluation, instruction, curriculum, supervision, staff development, budget, personnel, media: WHAT IS THE EXTENT AND KIND OF EXPERT FUNCTIONS NECESSARY TO SUCESSFULLY IMPLEMENT A DATA-BASED INSTRUCTIONAL IMPROVEMENT PROGRAM?

- how much and what kind of technical knowledge is available within district units or at the school site? how can those resources be tapped?
- if knowledge and expertise outside the district have to be utilized, where is it and how accessible is it?
- what factors strongly influence teachers' decisions about the content of instruction? How do these factors relate to the findings from testing programs?
- how might staff development activities deal with testing and evaluation especially with their linkages to instruction? Who would be prepared to conduct training in the development of appropriate tests, interpretation of the information they generate and the use of the information for diagnostic and prescriptive purposes?

- what effects does the current testing program have on the curriculum and how do these effects differ from the effects expected if a linkage system were established? How might this linkage shape the curriculum?
- what financial resources are available in the form of money, services or materials and what guidelines and procedures are required for their use? what strings are attached to the funds? what kind or support from the media can be counted upon and how would this support help the effort?

By coordinating mechanisms we mean both formal and informal structures that function to maximize staff commitment to, and staff communication about data-based instructional change. Our major question then would be: HOW CAN THE ORGANIZED UNITS AND LEVELS WITHIN THE DISTRICT AND THE ROLES OF DIFFERENT ACTORS BE CONNECTED SO AS TO PLACE TESTING AND EVALUATION ACTIVITIES IN A POSITION TO ASSIST TEACHERS? More specific questions might include.

- how might the information generated by testing and evaluation reach the different groups and in what form? Who would make use of this information?
- is there complexity in the organization of the district such that additional demands made by the change effort would be difficult to cope with? Would problems such as scheduling, limited personnel, staff turn-over reduce the likelihood of carrying the change through?
- what has the testing and evaluation branch of the district already done that has truly influenced instruction? How do the evaluation staff now relate to the staff of other branches and to the teaching

and administrative staffs at each school site? Does the evaluation branch staff have the skill to work effectively with diverse views, opinions and values?

- do different groups within the district communicate their meanings and intentions clearly, use appropriate decision-making methods and involve a wide range of appropriate persons in the decision-making process?

The last element in the framework, impact, includes the ways in which the management subsystem might affect instruction whether that effect is intended or unintended. The major questions then become: WHAT ARE THE DESIRED EFFECTS OF LINKING TESTING AND EVALUATION PROCEDURES TO THE INSTRUCTIONAL ACTIVITIES CONDUCTED AT THE SCHOOL? WHICH GROUPS DO WE HOPE TO IMPACT? WHAT MIGHT CONSTITUTE OBSERVABLE OUTCOMES? WHAT HARMFUL CONSEQUENCES SHOULD WE TRY TO ANTICIPATE AND AVOID? Specific questions might include:

- what instructional improvements would be considered satisfactory in relation to the effort made? How would this instructional improvement be measured? are there short and long range goals to be attained?
- which groups might experience impact as a result of the management strategy? How would impact differ from group to group? What would each group be expected to do as a result?
- how would ideas and attitudes related to testing and evaluation change as a result of the intervention?

Summary. The purpose of this paper was to suggest to educators contemplating a strategy for data based instructional change questions that might assist their thinking about the process. Data generated by tests and evaluation may be able to provide a sound basis for the management of the instructional system but such use requires a complex series of technologies and understandings.

References

Bank, A. "School District Management Strategies to Link Testing with Instruction." Presented to Evaluation Research Society, Austin, Texas, October, 1981.

Bank, A., & Williams, R. C. Projected framework for analyzing a school district's management subsystem for linking testing or evaluation with instruction (Technical Report #156). Los Angeles: Center for the Study of Evaluation, University of California, 1981.

Linking Testing and Evaluation Activities With Instruction: Can School Districts Make It Happen?

Richard C. Williams and Adrianne Bank

This paper was prepared for a symposium on NIE Funded Research on Testing and Evaluation in School Districts, American Educational Research Society, April 15, 1981, Los Angeles, California. It has been submitted to THRUST, a journal for district administrators.

The paper briefly describes some of the incentives and disincentives which districts experience in trying to manage data-based instructional change. What districts do appears to be quite varied. Here, one district's highly centralized approach is contrasted with another district's decentralized efforts.

Research into evaluation and test utilization from the point of view of the local school district trying to improve instruction is just beginning. The paper concludes with comments about the salutory implications of such research for practice.

Introduction

In July 1979, we began a three-year inquiry to discover ways in which school districts might effectively link their district testing and evaluation activities with instructional decision-making.

This inquiry was stimulated by our belief, based on previous research and experience in school districts, that testing and evaluation activities in most districts had only limited influence on internal school district instructional decision-making. Instead, the focus of testing and evaluation in many districts seemed to be toward satisfying external demands, e.g., federal program evaluation requirements, court-ordered desegregation mandates (Zucker, 1981; David, 1978). But many school districts had moved to develop their testing and evaluation capacities (Lyon, et al., 1978) and it seemed logical to us that the data and reports generated by a district evaluation unit might also serve as a district curriculum and instructional management information system.

The main purpose of our work is not to determine the extent to which a nationwide sample of school districts are using testing and evaluation for internal instructional decision-making. Instead, we are examining how a small number of districts are attempting to forge a linkage among testing and evaluation and instructional decision-making

At the present time, we have completed extensive case studies in five or six districts that we selected because they had a reputation for having tried to forge this linkage. Our sample districts, while not comprising a national sample, do exhibit characteristics that represent the diversity of American school districts. They reflect differences in:

size (large/small), student demographics (affluent/below-average income, racially homogeneous/racially heterogenious), and locale (urban/suburban). Three researchers have each spent approximately one week in each district visiting schools and district offices, interviewing district participants, examining relevant documents and records. We have asked respondants about three general areas: Why is this district trying to link testing and evaluation with instructional decision-making? How does this district do this? What effects have the linking activities had?

In the brief space available to us, we would like to discuss three specific questions related only to the first two areas of interest.

1. What are the incentives and disincentives that operate in school districts attempting to forge an evaluation-testing-instruction linkage?
2. What are examples of the approaches districts are taking to forge these linkages?
3. What are the potential contributions this research has for school improvement?

But before doing so, we'd like to define briefly what we mean by linkage although you will get its fuller flavor, by example, later in the paper. Linkage, to us, means the coordination--either through formal or informal means--of all the operations and services within a school district essential or supportive of the use of testing and evaluation for instructional purposes. Linkage is a function of management. It is an arrangement which brings together in some productive manner data collection, analysis and reporting with core instructional activities.

Such testing-evaluation-instruction linkages are not commonplace in school districts although testing and evaluation activities have increased substantially since 1965. This may mean that most school districts have, over the past 15 years, felt little need to make such a linkage. We were interested to learn what factors seemed to be encouraging our sample districts to move in this direction.

Question 1. What are the incentives and disincentives that operate in school districts attempting to forge an evaluation-testing-instruction linkage?

In the districts we studied, the single shared reason given for initiating coordination arrangements between tests and evaluations was to influence pupil achievement as measured on test scores. In many of the districts there had been expressed dissatisfaction, coming from a number of sources, with the academic performance of students. The move towards use of tests and evaluation data was primarily remedial. In one of the districts, however, there had been overall satisfaction with student learning; moreover, there was a sense, on the part of the district superintendent, that individualized instruction might increase the learning of average and above average students.

District officials indicated in their interviews with us that their overall intention was to use test scores as a description of student achievement. They wanted these scores arranged and understood in such a way so as to redirect instruction. However, the immediate incentives for starting and continuing such a process seemed to vary from district to district. For example, some central offices were moved in this direction by explicit mandate from courts, or from state legislatures or from school boards. In other districts, superintendents or other officials seemingly influenced by research and current educational thinking, decided to use available federal and state money to build instructionally relevant tests.

We might categorize the types of incentives we found as either "sticks" or "carrots" and their sources as either external or internal to district management. Our matrix would look something like this.

---

insert Figure 1 about here

This list of incentives, to some extent, begs the question. The carrots and the sticks are common to other districts. Why haven't they moved to link testing and evaluation with instruction? Given our small sample, and our field-based research design, we cannot provide a general answer to that question. What we can say is that certain characteristics seem to be present in our five districts, especially those that are most advanced in thier linkage development. These elements indicate that our districts had the management capacity to respond to the incentives. The elements we refer to are: idea champions, stable core staff, realistic problem analysis, and tolerance for ambiguity. The following is a brief description of each element:

- Idea champions--by this we mean individuals in key administrative and policy positions who firmly believe in the value of test and evaluation data and consistently champion its development and connection to instruction. In our districts, these individuals were found in a variety of positions. There was no consistent pattern to their school district assignments, e.g., some are in curriculum, some in evaluation, some are line administrators; what they do share with

	Sticks	Carrots
External	•requirements by federal or state agencies to: •evaluate problems •develop courses of study •raise test scores •community dissatisfaction with public education expressed by: •press and media •loss of students •Board action	•availability of federal and state money for •text development •evaluation of programs •staff development •relationships with universities •existence of techniques or procedures to link tests with instruction
Internal	•decision of district administrators to link testing, evaluations and instruction	•desire of district to acquire additional funds •presumed likelihood of success in linking testing and evaluation with education

Figure 1

one another are characteristics such as conviction, persuasiveness and some degree of power.

- Stable core group--in our districts, these "idea champions" and their followers have been around for a while. In our most advanced linkage systems it has taken from 8-10 years for the linkage programs to develop and mature. This could not have happened if the core group had continually changed.
- Comprehensive rather than ad hoc problem analysis--the core group has been aware of next steps beyond the immediate task of the moment. It is one thing to develop, for example, a CRT program in reading--it is quite another thing to actually get teachers to use it. Bridging the gap between development and use implies an understanding of the school site and district as a bureaucratic social system and an appreciation of the various strategies and tools that might most effectively bridge the gap.
- Tolerance for ambiguity--none of the linkage arrangements developed, over time, in anything resembling the rational, linear way that is often described in standard planning texts. Instead, the programs have developed unevenly, component by component on a broken front. Many times, the components of the linking system have been developed independently of one another, with different purposes and each with its own set of advocates. Developing linkage arrangements to merge together these disparate pieces into a new configuration takes time and it can be very frustrating.

The above list is not exhaustive and it may be that these characteristics and activities are found in many districts that have not thought about or who have tried and abandoned an effort to develop a linkage system; we don't know. All we can say is that these are present in our sample districts and we believe they contribute to the progress these districts have made.

What about disincentives to develop a management arrangement that links services and supports to the connection between testing, evaluation and instruction? As we indicated above, typically districts are not moving in this direction. There are likely several reasons for this. An important one, we believe, is that these districts are not pressured or pulled to think about the impact of students' test scores through change in instructional activities. Thus, they continue in a traditional arrangement of semi-autonomous operational units.

For example, districts may feel that their declining test scores are caused by large and rapid changes in the ethnic or racial class make-up of their pupil population or that their declining performance merely reflects the flagging public support for the schools. They reason that, until these conditions change it is unlikely that encouraging curriculum and instructional changes based on test scores and evaluation findings will make much of a difference. They conclude that other political, social or financial strategies might be more appropriate.

Another disincentive is that closely linked testing, evaluation and instructional system, with its emphasis on supervision, communication, and coordination, flies in the face of the traditional school district operating mode which can be characterized as loosely coupled, (Weick, 1976) with teachers working quite independently behind closed classroom doors (Lortie, 1975). Teachers do not readily embrace approaches that fundamentally alter their accustomed professional behavior patterns.

Still another disincentive may be that a tight and interactive relationship between test scores and classroom practice is yet an unproven solution to the problem of student learning. While various components, e.g., development and use of CRTs, formative and summative evaluation methods, are becoming increasingly sophisticated and technically sound, much remains to be done before teachers and administrators are convinced that these techniques can be used as effective tools in their own classrooms for improving student achievement. Given the other demands on their time and energy, teachers will not readily commit themselves to unknown and unproven technologies.

Question 2. What are examples of the approaches districts are taking to forge these linkages?

Our sample districts are using a variety of approaches in linking testing and evaluation with instructional improvement. In this paper we will very briefly describe two approaches: a decentralized, school-oriented system using norm-referenced standardized test scores; and a district-directed centralized system using district-directed criterion-referenced tests. The decentralized NRT system uses the individual school as the locus of change. Within loosely prescribed district parameters, each school has considerable discretion in developing and implementing an instructional program that the school staff feels is appropriate for its particular student body. The norm-referenced student test results are folded into an individualized evaluation report that is prepared for each school by the district office. The school staff, with the assistance of the central evaluation unit staff and, often, with district instructional and curricular specialists, develop yearly plans in which they identify its own instructional and other programmatic priorities. Presumably, the evaluation unit's reports, including the scores, form part of the evidence upon which each individual school modifies its instructional program. Some of these districts were also developing and using CRTs, but these tests did not play a prominent part in their instructional renewal program; they were used more as an instructional tool in the classroom rather than as a tool for school-site decision-making.

The school districts using a centralized CRT system focus on a common district instructional continuum, usually in reading, math and language arts to which all schools are expected to adhere. The impetus for change comes more from the district level, than from the local school. The district also encourages the teaching staff to follow a common instructional methodology when implementing the district's curriculum. Student scores on CRTs are used as the main basis upon which instructional effectiveness is gauged. The CRTs are developed so that they relate to the district's adopted instructional program. NRTs are administered and reviewed but they are used mainly to inform the public of the district's program--they do not play a prominent part in the instructional renewal program.

We do not wish to imply that NRTs are not appropriate for decentralized systems or that CRTs are inappropriate for decentralized systems. We are merely reporting that these were the configurations we observed in our small sample of districts. Likely other mixtures of these elements have been devised.

Although the decentralized and centralized orientations differ in the locus of change and the types of test that are used, the districts' arrangements share important characteristics such as providing support services to the schools: e.g., an extensive and appropriate in-service component, a well-developed data processing capability, a skilled evaluator and measurement staff.

The districts differ in regard to what they considered the effect of their programs. The two centralized, CRT-system districts pointed to what they consdiered substantial improvements in pupil achievement as a result of their program. The decentralized districts were less sure of the overall effect of their program on student achievement but cited process changes at the school in evidence of effect. This is understandable since the schools themselves differ in what they are trying to accomplish; and these diverse intentions do not lend themselves to more standard yardsticks of progress. Of course, it may be that it takes longer to see the effects of a decentralized program than a more centralized one. We are not yet prepared to offer reasons for, or to assess the differences in the effectiveness of the two approaches. The districts themselves were not presently examining what might be considered unintended or unexpected side effects, e.g., heightened or lowered teacher morale, increased or decreased community support.

Question 3. What are the potential contributions this research has for school improvement?

There is a substantial public and professional "crisis of confidence" in the public schools' ability to adequately educate its pupils--

especially in the basic skills. Increasingly, districts are realizing that methods of school improvement built on piecemeal approaches, such as untargeted in-service training programs, or new testing programs, or adopting and implementing externally funded projects. These activities, however well-intentioned, simply were not reversing the declining test scores.

Some districts, such as our sample districts are now seeking more comprehensive and integrated approaches to developing better teaching and learning. One such approach involved connecting the school districts' testing and evaluation activities with on-going discussions about how to chart the district pupils' achievement, assess the effects of various instructional strategies, revise those strategies and use subsequent data to re-assess. We believe this systematic approach will be increasingly tried by other districts. While we think that each district will have to evolve an approach that is appropriate to its particular context and needs, it seems logical that districts beginning to consider this approach can learn a great deal from the experience of these "pioneer" districts. They can learn of the various strategies that have been tried, the specific components (such as CRTs) that have been developed, and the kinds of barriers that have been encountered. Enlightened by the experience of those who have preceded them, these "newer" districts can, perhaps, reduce the time and cost necessary to implement such a system.

Our sample districts have been deeply involved in developing these programs and this has made it difficult for them to step back and take a comprehensive and somewhat detached view of their efforts. What is more, they do not have the opportunity to compare their efforts with those of other districts that are developing a similar linkage strategy.

We see ourselves as providing two research-related services; first, as observers and recorders of what these districts are doing, so as to subsequently create from their synthesized experiences technical assistance materials for districts wishing to follow this linkage strategy as a means of improving pupil achievement; and second, as analysts of this process, we seek to understand the configuration of human, organizational, political, and technical elements that are associated with the implementation of this linkage strategy so as to contribute to the growing school improvement literature.

With regard to our technical assistance and development role, we realize that the linkage arrangements that our sample districts are developing are unique to each setting and that they cannot be "packaged" and exported to other districts. Nonetheless, there are likely portions of these arrangements that can provide guidance to other districts. The things these sample districts have learned about the process will likely be of considerable interest to those who want to embark on this strategic course. During the last year of this project, we will be working with several districts and helping them begin to design and implement such a program.

With regard to our research/analytical role, we see as a major contribution the bringing together of the research literatures from several fields, e.g., evaluation, testing and curriculum and organizational theory as a means of gaining insights into the dynamics of this linking process in school districts. Since these literatures have historically been developed in isolation from each other, our research provides a unique vehicle for gaining a better understanding of their interrelationships. This kind of theory/practice synthesis seems to us to be a necessary step if we are going to be able to fashion research and conceptual work into tools useful for working on the pressing problems facing public education today.

Bibliography

David, J. Local uses of Title I evaluation. Palo Alto, California: Stanford Educational Policy Research Center, 1978.

Lortie, D. C. School teacher. Chicago, Ill.: University of Chicago Press, 1975.

Lyon, C. D., Doscher, L., McGranahan, P., and Williams, R. Evaluation and school districts. Los Angeles: Center for the Study of Evaluation, UCLA Graduate School of Education, 1978.

Weick, K. E. Educational organizations as loosely coupled systems. Administrative Science Quarterly. 1976, 21, 1-19.

Zucker, L. G. Institutional structure and organizational processes: the role of evaluation units in schools. In Evaluation in school districts: Organizational perspectives (CSE Monograph #10). A. Bank & R. C. Williams (eds.). Los Angeles: University of California (in press).

How Do External Factors Influence School District Management: A Preliminary Inquiry

This paper is an exploration of one of the elements in the Framework--that called relevant environments. After considering the work of Pfeffer and Salancik in The External Control of Organizations - A Resource Dependence Perspective, we describe the different environments of two school districts. We then speculate on the connection between environmental factors and the degree to which these districts have centralized their management of instruction.

HOW DO EXTERNAL FACTORS INFLUENCE SCHOOL DISTRICT MANAGEMENT: A PRELIMINARY INQUIRY

Richard C. Williams
Michele Marcus

Introduction

For the past two years project staff at UCLA's Center for the Study of Evaluation (CSE) has been studying ways in which school districts can effectively link their district-wide testing and evaluation activities with district instructional programs. Previous research (Lyons et al., 1978) had convinced us that most school districts had not forged such a linkage; testing and evaluation had remained largely uncoupled from the central instructional program (Meyer & Rowan, 1977; Williams, 1979).

Based on recommendations from knowledgeable colleagues in the research and practitioner communities, we identified six school districts that had reputations of having made exemplary efforts to link their testing and evaluation efforts to their instructional core. Subsequently we conducted case studies in these six districts to see whether or not they had forged such a linkage, to determine the processes and structures they had employed and to see if there were any generalizations and insights we could derive from these districts' activities that might be useful to other districts wanting to forge similar linkages. We found, not surprisingly, that the districts differed in the progress they have made in their program and in the structures and processes they have been using (Williams and Bank, 1981). After having described these programs in some detail we have now begun trying to understand the variations we have observed. For example, some districts have developed a district-wide plan based upon a common district instructional continuum. That district is developing a common set of expected teacher competencies and behaviors and the testing and evaluation programs are tightly coupled to that emerging technical core. We refer to this as a centralized approach.

Other districts, in contrast, have a much more decentralized approach in which the unit of change seems to be the local school site. While school site data may be collated and compared district-wide, each school site is considered the main unit of analysis and change;and the testing and evaluation programs are more loosely coupled to the instructional core through intermediating local school sites. We refer to this as a decentralized approach.

What accounts for these two different approaches? Both seem to be, or have the potential to be, successful. No doubt the different approaches have resulted from both carefully considered as well as accidental factors. That is, those who have been involved in the development of the systems likely had some preconceptions about the advantages of centralized versus decentralized approaches. Administrators supportive of one or the other position could probably marshall arguments from the organizational theory literature to support each of their views. However, educational organizations, in common with other organizations, experience twists and turns in directions due to the arrival or departure of key actors at critical times.

The belief systems that influence organizational design and the historical condition that surrounds such decisions have been recognized for many years by researchers and practitioners alike. However, much of their attention, we would argue, has been directed towards the internal workings of an organization. Who are the powerful organizational leaders and policy makers? What belief systems guided their thinking? How can the organization's design be made most compatible with the organization's personnel? What internal coordinating or authority system will work best given the organization's personnel and design? Certainly these internal organizational characteristics and conditions are essential to designing and implementing a decentralized or centralized system but all suggest that even deeper insight can be gained when one considers, in addition to the internal factors, the external factors.

All organizations exist within a number of relevant environments and have interactions with them. Organizational boundaries are penetrable by outside influences. This permeability means that organizations cannot function isolated from such external factors as funding sources, client characteristics and preferences, legal and legislative mandates, and unexpected events such as floods, recessions and population shifts.

Organizations differ with regard to their boundary permeability. Public school districts, with their publically elected school boards, high client interest in pupil performance, and public control of funding, represent highly permeable organizations. It follows that the governing and operating structures of public school districts are likely to be influenced by external factors--and that a better understanding of public school district design and functioning can be understood when both internal and external factors are examined. More specifically to the topic of decentralized and centralized testing and evaluation and instructional subsystems, is whether there is some relationship between the extra organizational conditions and decisions to use a centralized or decentralized approach. Or, can one better estimate a selected approach's chances of success within a given district when only internal conditions and factors are considered.

Our purpose here is to speculate a bit on the influence of external environmental factors on centralized and decentralized testing, evaluation and instructional subsystems. We do not suggest that this examination of external factors will result in any hard and fast set of rules that will settle conclusively the merits of one approach over the other. Instead we want to raise the "consciousness level" of those who work within such systems so that they consider both internal and external organizational factors. In the following paper we will:

- discuss briefly the theoretical perspective that guides the consideration of external factors upon organizational design and processes;
- describe case studies of two districts--one using a centralized and the other a decentralized linking subsystem--and focus on the role and influence the external environment has on the centralization-decentralization approaches;
- discuss implications these observations have for those considering a centralized versus decentralized approach to linking testing or evaluation with instruction.

A Theoretical Perspective On External Influences On Organizational Structures and Processes

A major step in understanding organizational functioning was the adoption of general systems theory (Bertalniffy, 1937; Katz and Kahn, 1966) when analyzing organizations. Prior to using general systems theory, organizational analysts had focused on internal matters and had largely ignored the role and function an organization's external environment may have had. But general systems theory properly placed organizations in the perspective of a functioning unit that has continuous interaction across its boundaries--influencing its environment and being influenced in return by the environment.

A number of theorists have speculated on and conducted research on that phenomenon and its influences on organizational functioning. One of the earliest speculations on this phenomenon was that of Burns and Stalker (1968) whose research on the post-war electronic firms identified mechanistic and organiz organizations.
Lawrence and Lorsch (1967), for example, have, on the basis of research, evolved a contingency theory which seeks to explain how the number of components in the environment and their characteristics can, or should, help determine an organization's function and design. That is, an organization's function and design should properly be flexibly contingent upon the external environment's characteristics. Derr and Gabarro (1972) applied that work in analyzing the Boston Public Schools.

Perhaps the major recent influential work on this topic has been that of Pfeffer and Salancik (1978), who develop a "model of environmental effects" which can be applied to both private sector and public sector organizations. Their main thesis is that external environmental influences exert control on the internal workings of an organization and consequently help shape the organization. They contend that "to understand the behavior of an organization you must understand the context of that behavior--that is, the ecology of the organization (1978:1)." Part of the problem in understanding the environment is that the environment of an organization can affect an organization's outcomes without affecting its behaviors. This occurs because important elements of the environment may be invisible to organizational decision makers, and therefore, not considered by them in their shaping of organizational actions; but these same elements, independent of administrators' perceptions, do affect organizational success or failure. For example, in the early 1960's when some American firms decided to purchase coal mines, it is doubtful that they gave much thought to the Arab world when making these investments. In the 1970's, however, when Arab governments raised oil prices, many of those companies who had invested in coal profited. Outcomes were affected by external events even though it is unlikely that the original decisions had been influenced by them.

Pfeffer and Salancik present the model by which the environment is linked to organizational change and action. The model suggests that the relationship between environments and organizations is not random but is indeterminate, and that the very indeterminacy of environmental effects on organizations is potentially explainable. As an example, the model plots the effects on the organization of executive succession--the removal of one executive and the selection of another. The authors contend that both the removal and subsequent selection of top administrators is affected by the organization's environmental context.

Pfeffer and Salancik's model of organizational change can be summarized briefly: (1) the environmental context--with its contingencies, uncertainties, and interdependencies--influences the distribution of power and control within the organization; (2) the distribution of power and control within the organization affects the tenure and selection of major organizational administrators; (3) organizational policies and structures are results of decisions affected by the distribution of power and control; and (4) administrators who control organizational activities affect those activities and resultant structures. Executives are a source of control, and it matters who is in control because control determines organizational activities. The environment affects organizational activities because it affects the distribution of control within the organization (Pfeffer and Salancik, 1978:228).

Pfeffer and Salancik use this model to highlight three seemingly causal linkages that may connect environmental factors to organizational characteristics. First, a link exists between the environment--a source of uncertainty and constraint--and the distribution of power and control within the organization. Second, a link exists between the distribution of power and control and the choice of executives and their tenure. Third, a relationship exists between organizational executives and the actions and structure of the organization. One may not observe a perfect relationship among these links because, according to Pfeffer and Salancik, organizations are only loosely coupled with their environments, and power is only one important variable intervening between environments and organizations.

External Environmental Influences - Two School Districts

Admittedly, the ways in which organizational structure and behavior are constrained by forces in the environment are different for different types of organizations. Private sector organizations which focus on producing and delivering goods may be affected by the buying trends of the public whereas public sector organizations which are concerned with delivering services may not be influenced at all by sales or marketing trends. Industrial organizations which fail to take environmental variables into account when making strategic decisions risk losing their competitiveness in the market place. Each industry depends on the demand for its products to maintain its supply of customers and revenue, and thus its very survival.

At first glance the public sector, especially the public schools, would seem to be more environmentally free than industry. Carlson (1964) has referred to public schools as disinterested organizations which are guaranteed their resources and clientele. This has the effect of diminishing the public schools' resolve to respond to external environmental influences and the pace and adequacy of response to environmental changes is comparatively weak. Likely this phenomenon is true but this should not blind school administrators and analysts to the effects the internal environment can have and the symbiotic relationship between school district structures and process and the external environment. Zucker (1981), for example, has argued that because school districts are "institutional" rather than "technical" organizations, they must perform in accordance with public prescriptions and expectations rather than attending primarily and exclusively to their technical--i.e., instructional--functions. When school administrators are reflecting on their own structures and functions they should consider environmental conditions and characteristics.

As a means of illustrating some of these, and for purposes of helping those who wish to derive an appropriate organization configuraion for linking testing, evaluation and instruction, we turn next to two case studies of districts that have derived a testing, evaluation and instruction subsystem. One is centralized; the other decentralized. We will speculate on how external environmental considerations have shaped the structures and activities being used and we will discuss the "fit" between external conditions and each school district's approach.

Two Case Studies

We have selected two districts--one using a centralized approach (Crescent City) and the other using a decentralized approach (Bordertown). From an admittedly large number of external conditions we have selected the following three characteristics:

- Population mobility;
- External mandates;
- Religious and cultural conditions.

Using the Pfeffer and Salancik conceptualization we will link those characteristics to executive succession. Finally we will discuss implications that this approach has for understanding organizational functioning.

Population Mobility. Crescent City School District is an urban-rural district with a 79 percent Anglo population, experiencing a surge of growth in its student enrollment. Since 1970 the district has added 17,000 pupils. As a result the district has built new school buildings and hired more teachers. In addition, the city's major industry encourages considerable population mobility. Many families come and go regularly and there is considerable movement among school attendance zones. In order to provide some consistent educational program for pupils who move from school to school the district has abandoned its somewhat decentralized approach to curriculum and instruction and has adopted what many districts would consider a very centralized approach.

Bordertown School District, on the other hand, is experiencing a decline in student population. Between the 1964-65 school year and the 1976-77 school year the district's enrollment declined by 22,500 pupils. As a large urban school district, it is experiencing "white flight" and is witnessing a slight influx of black students annually. Currently, 56 percent of the district's pupils are black. Moreover, a small percentage of minority students from a neighboring state is moving into Bordertown. These students are characteristically poor, unschooled, and illiterate; the parents are extremely protective of the children and suspicious of the schools. Even though there is some transiency both into and within the school district, more students are exiting than entering. Too, the heterogeneous quality of the population is an environmental constraint against any mass mobilization effort to centralize the schools. Consequently, for this and for cultural reasons to be discussed in a future section, Bordertown has adopted a relatively decentralized approach to school district curriculum and instructional management.

External Mandates. The Crescent City School District has pro- programs (e.g., ESEA Title I, State Minimum Competency Testing) many of which have state or federally mandated evaluations. The District is obligated by law to comply with such policies. The State and/or Federal government also provides an increasing percentage of the District's budget. The state is currently controlled by a fiscally conservative governor and legislature and is subject to reductions in financial support. Although the District already has a low expenditure per pupil, ranking near the bottom nationally, more budget cuts are planned. With less money allocated to schools, the District is operating under considerable financial restraints. A result of this has been an increasing level of internal conflict between the organized teachers and the school board and administration over salaries and working conditions. This has influenced teacher attitudes toward the administration and played an important part in the school superintendent's recent resignation from his post.

Although it is not controlled by State minimum competency testing mandates, Bordertown must comply with a State mandated "graded course of study." This governs the scope and sequence of subjects taught in the public schools. The legislature has also recently reduced the funding allocations for urban public schools; this political body has a reputation as being a "pro-suburb advocate," and many District officials feel that it neglects the urban areas and their problems. Bordertown, however, does receive additional funding through ESEA Title I and Title IVC programs which have allowed the District to create and implement some innovative programs of its own. In addition, Bordertown School District receives extensive funding through Federal vocational education sources. In fact, approximately 50% of its secondary pupils are enrolled in vocational education programs. An additional evnironmental constraint is apparent in the State mandate that vocational education teachers must also teach regular subjects (e.g. English, Math, etc.). Located in a large manufacturing center, the District's vocational education programs receive strong support from the business community.

Other external organizations are influential in bringing about changes in the District. For example, the teachers' union has had sucessful strikes in the past, and still exerts pressure on district decision makers. Community groups too form coalitions for particular causes and exert pressure on the District's administration. For example, the existence and power of community task forces changed the District's procedures for evaluating its alternative schools.

Religious and Cultural Conditions. Although Crescent City was founded in the mid-1800's, it remained a tiny watering spot on the westbound trail until after World War II. In the early 1950s a growth spurt began and today it is one of the larger cities in the nation. Still, it is a relatively young town with a somewhat homogeneous population. Most of the District administrators now in top level positions immigrated to Crescent City in the mid 1960s. Consequently, the "traditional way" of doing something was non-existent. Attention to current commitments is more characteristic of the District's leadership.

Crescent City is now the largest city in the State and the District educates 59% of the State's pupils. The city is surrounded by desolate areas with small rural communities as its only neighbors. Therefore, the District represents an educational monopoly; there are virtually no competitive public or private schools to drain off pupils or to attract teachers. The large Mormon population promotes the separation of church and State and the separation of family responsibilities from school responsibilities. Thus, benign support is given to District policies unless they interfere with family responsibilities (e.g. sex education) or fall short of expected performance levels (e.g. student test scores). For example, there was no public outcry when the District recently instituted an attendance policy requiring failing grades to any student absent more than a set number of days; instead community members accepted, and indeed supported, the policy.

Bordertown is a densely populated area with many suburbs and other major metropolitan cities nearby. Approximately one fourth of the school age children attend private or parochial schools. The large Catholic population staunchly supports the Catholic schools. Thus, Bordertown School District faces tough competition in attracting high-achieving students and quality teachers. Community members often compare--unfavorably--the public schools to the private schools. The public school officials complain about the unfairness of the criticism considering the constraints the public schools face in acceptance of clients and availability of resources.

Moreover, Bordertown lives with a strong sense of history. Founded in 1788, it was the nation's sixth largest city and third largest manufacturing center by 1860. There are many stable, old neighborhoods whose natives wouldn't conceive of doing anything which would violate Bordertown's past culture. In fact, Bordertown has been called a "city of cities" where these neighborhoods are identifiable by race, ethnicity, and social class. Consequently, decision-makers are often tied to tradition and fearful of untested solutions to local problems. In each neighborhood active community task forces, or "forums," serve to protect local interests on matters such as zoning, road construction, and schools. The diversity of the population has led Bordertown School District to adopt a decentralized approach to education and to establish many types of alternative schools.

As Pfeffer and Salancik (1978) contend, the environmental context, with its contingencies, uncertainties and interdependencies, influences the distribution of power and control within the organizations. Then the distribution of power and control within the organization affects the tenure and selection of the major organizational administrators. Finally the organizational policies and structures are results of the decisions affected by the distribution of power and control. Admittedly, the administrators who control organizational activities affect those activities and resultant structures. The histories of executive succession to the Superintendency for both Crescent City and Bordertown serve to illustrate Pfeffer and Salancik's "model of environmental effects" regarding executive succession and organizational change.

During the heyday of change and innovation in the 1960s Crescent City's Superintendent emphasized local school building autonomy--each school was to develop its own program tailored to its pupils' needs. Following through with the administration's decentralized approach, the District was subdivided into four administrative zones with considerable autonomy in each zone. When that Superintendent resigned to accept a superintendency position with another district, he was promptly replaced with an administrator who shared his views and would continue his policies.

In the late 1960s and early 1970s, environmental conditions changed. The community became concerned over pupils' low test scores, desegregation, and the educational inequalities of Crescent City's decentralized system. Thus, that Superintendent was forced to resign by pressure from the community and the Board of Education. A new superintendent who would address the current issues of concern was appointed. This Superintendent guided the District through desegregation and began the centralization process by eliminating the four-area decentralization scheme. He appointed one deputy and four associate superintendents who ran the District's central administration. He also allowed certain administrators to begin revising and centralizing the District's instructional program. When this Superintendent chose to resign to enter the private sector, a successor who was committed to a centralized curriculum was selected. More recently a crisis between the teachers' union and the board precipitated by limited district financial resources, played an important part in this Superintendent's decision to resign.

Therefore, in each case of executive succession the selection of the new superintendent seemed to be a reflection of the environmental context, which in turn influenced the distribution of power and control within the District.

The minority population of Bordertown was concerned with desegregation in the early 1960s. In 1963 the Board of Education successfully defended a desegregation suit brought by the NAACP. Although the federal district court found, and the Court of Appeals affirmed, that no alleged discriminatory practice on the part of the Board brought about the racial imbalance that existed, many community members were dissatisfied with the school district's policies. Neighborhood associations exerted pressure on the Board of Education to reduce the racial isolation of Bordertown's schools. Consequently, in the early 1970s the Board hired a new liberal superintendent who favored integration and had a successful record for integrating schools and implementing innovative programs. The new Superintendent instituted an administrative decentralization plan, creating six area directors. He then promoted a number of principals (including several black principals) to these new positions, thereby installing a new echelon of administrators loyal to him. In addition, he was influential in getting the Board to adopt a policy establishing integration as a high District priority, and also in establishing an open enrollment policy which allowed students to attend any District school with available space providing the transfers would improve the racial balance. The administration also began plans for the city's first two alternative schools. By the mid-1970s the environmental conditions had changed and a more conservative Board was elected. The Superintendent resigned under pressure from the Board and a more conservative Superintendent succeeded him.

Thus, as in Crescent City, the removal of the Superintendent and the naming of the successor seem to be reflections of the environmental influences upon the balances of power in the District.

Implications of the Findings

In reviewing the literature many authors point out that understanding the relevant environments is important for understanding organizational actions and structures. In the past many organizations seeking to increase their effectiveness have adopted other organizational patterns, policies, and/or strategies on the basis of internal conditions without considering the external conditions. Pfeffer and Salancik contend that external environmental influences exert control on the internal workings of an organization and help to shape the organization. Admittedly, our research seems to suggest that organizational patterns, policies, and strategies are indeed reflective of the external environmental conditions encompassing the organization.

Thus, school district administrators wanting to implement some organizational change need to understand the ecology of the organization, the environmental context of the behavior of the school district. By addressing the external environmental conditions as well as the internal organizational conditions, administrators can select and implement successful change strategies. By considering all of the relevant variables--population mobility, pressures from special interest groups, available resources to name a few--an optimal system for increasing organizational effectiveness could be developed.

References

Aldrich, H. E. Organizations and environments. New Jersey: Prentice-Hall, Inc., 1979.

von Bertalanffy, L. General systems theory: A new approach to the unity of science. Human Biology, Vol. 23, Dec., 1951. pp. 302-361.

Burns, T., & Stalker, G. M. The management of innovation. London: Tavistock, 1968.

Derr, C. B., & Gabarro, J. J. An organizational contingency theory for education. Educational Administration Quarterly (8), 1972, p. 26-43.

Carlson, R. O. Environmental constraints and organizational consequences: The public school and its clients." N.S.S. Yearbook. University of Chicago Press, 1964.

Hawley, A. H. Human ecology. New York: Ronald Press, 1950.

Katz, D., & Kahn, R. L. The social psychology of organizations (1st ed.). New York: Wiley & Sons, 1966.

Lawrence, P. R. (ed.). Studies in organizational design. Illinois: Irwin and Dorsey, 1970.

Lawrence, P. R., & Lorsch, J. W. Organization and environment: Managing differentiation and integration. Cambridge: Harvard Graduate School of Business Administration, 1967.

Lyon, C., Doscher, L., McGranahan, P., & Williams, R. C. Evaluation and school districts. Los Angeles: Center for the Study of Evaluation, University of California, 1978.

Meyer, J., & Rowan, B. Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, 1977, 83, 304-363.

Pfeffer, J., & Salancik, G. R. The external control of organizations: A resource dependence perspective. New York: Harper and Row, 1978.

Reich, R. B. Industries in distress. In The New Republic, 184:19-20+, May 9, 1981.

Williams, R. C. "School District Characteristics and Their Impact on Evaluation Offices." Presented at AERA, San Francisco, April, 1979.

Williams, R. C., & Bank, A. "Linking Testing and Evaluation Activities with Instruction: Can School Districts Make It Happen?" Presented at AERA, Los Angeles, California, April, 1981.

Zucker, L. G. Institutional structure and organizational processes: The role of evaluation units in schools. In Evaluation in school districts: Organizational perspectives (CSE Monograph #10). A. Bank & R. C. Williams (eds.). Los Angeles: University of California (in press).

School District Management Strategies to Link Testing with Instructional Change

This paper was presented at the Evaluation Research Society, Austin, Texas, October 1981. It will be revised and submitted for publication either in a journal of general educational interest or one read by educational administrators.

In the paper, we describe important pre-conditions for linking testing with instruction using some of the elements outlined in the Framework. We then describe three management strategies which we observed within our six districts. Although the point will be made more sharply in a future version of the paper, these strategies may be handled either in a decentralized or in a centralized manner, as alluded to in the preceding papers.

School District Management Strategies to Link Testing with Instructional Change

Adrianne Bank, Ph.D.
Center for the Study of Evaluation

Summary: Based on two years of field work in six school districts, I discuss in this paper some conditions which appear to be present in school districts who are trying to manage the dissemination and use of test data for instructional change. I then describe three strategies--or, more precisely, three configurations of linkage activities--which we saw in our districts. Finally, I list several characteristics which these strategies have in common, and suggest some implications for school districts interested in the management problems associated with connecting testing with instructional change.

Introduction

This paper is a trial balloon. It is a think-piece in which I am formulating ideas that have evolved from two years of field experience in school districts and schools as well as from continuous dialogue among CSE project staff. Egon Guba in his CSE monograph (Guba, E., 1978) notes that naturalistic research occurs in expansionist and reductionist waves, waves which alternate between discovering data and making sense of data. Here and now, with this wave, we are reducing and making sense of the data. Your comments and questions will help in this process.

Let me briefly summarize how we "discovered" the data. We selected six school districts, through an extensive nomination process, who had a reputation of "doing something interesting" to link their testing or evaluation activities with instruction. Our teams then spent several person-weeks in each district talking with and interviewing as many as forty individuals per district, from Board members and parents to principals, teachers and students. We started by trying to understand what these individuals thought they and their district were doing in relation to testing or evaluation and instruction. In other papers, we discuss what we learned about how the districts came to be doing what they are doing, and what impact their policies appear to be having. Here, we will only try to make sense of what they say they are doing, or what we observed them to be doing, to link testing with instructional change.

And, even within this narrowed framework, we are putting aside a whole range of important technical questions, among them the quality of the tests themselves, the initial match between the test and the instruction, the procedures used to analyze test responses. All of these are essential to connect testing programs with ongoing instruction, but for the moment, we are relegating them to the sidelines in favor of discussing management concerns. And, we are not even addressing what we regard as important up-front management issues--such as the involvement of the community and teachers in the selection or development of the testing instruments. We are today looking at only the "back-end" of the testing process--the dissemination and use of testing for instructional change.

Just a word about definitions. By testing we refer to district-wide testing programs in which specified populations of students respond on norm-referenced achievement tests or criterion-referenced diagnostic tests or proficiency examinations or state assessment programs. We are excluding from this discussion teacher developed quizzes or unit tests embedded in curriculum materials. And, by data-based instructional change, we mean any decision or activity resulting either from the tests themselves or from an analysis of students' test scores that alters the way in which teachers and children in classrooms feel, think or behave.

Before proceeding, I want to make three brief preliminary comments about the concerns and assumptions that Dick Williams, my co-director on this project, and myself bring to our research effort. First, we are interested in school district administration in the area of testing, evaluation and instruction and the effect of district policies and procedures on schools and classrooms within the district's preview. The activities and attitudes of principals and teachers are sometimes viewed in isolation from district influence--as if these individuals' work environment was bounded by the building itself and the attendance area of their students. One aspect of our work, therefore, is to explore the extent to which district management of tests and of instruction and of the link between them--good or bad, strong or weak, loosely or tightly coupled--impacts schools and classrooms. Second, we know that, over the past fifteen years, awareness about tests and the capacity to use them have been developed in many district central offices. We are curious as to whether that capacity can be turned from satisfying outsiders'--that is, federal and state legislators or administrators--demand for information to stimulating insiders--that is, administrators, principals and teachers--towards instructional change. Lastly, and most importantly, we are more concerned with instruction than with testing, and we recognize that it is likely that there are other, more potent instructional improvement levers around. Nonetheless, since testing is ubiquitous in American schools and since test scores seem to have the potential to catalyze instructional change, we have been investigating how this might happen.

Conditions Needed to Link Testing With Instructional Change

Let us turn to our six districts--one small, two medium and three large--and report some of what we found.

- We found all six districts in the midst of tumult and problems. Other researchers, such as Mary Kennedy in Huron's recent study of 18 districts found similar situations (1980). Among the crises effecting our districts were: court interventions into both desegregation and instructional matters; massive population shifts requiring the closing of some schools, possibly the opening of others; large numbers of children with first languages other than English; budget cuts; low teacher morale; vociferous and divided community opinions on the goals schools should emphasize, etc.;

  We found four district where instructional programs, testing programs or evaluation cycles had been developed in response to the availability of federal or state fundings. Many people in these districts had spent time writing grant proposals and reorganizing their operations to meet categorical program requirements. However, in two other districts, administrators had resisted the federal temptation and made do with general funds;

  We found, for the most part, capable people in district offices, in schools and in classrooms, doing their jobs and concerned about children's learning. Within and across districts there were some teachers, principals, administrators and parents who were frustrated at problems they saw as beyond their control, others who had a sense of purpose and saw the possibility of improvement;

  We found central office organization charts which defined district operations such as curriculum, instruction, testing and evaluation, personnel, budgeting, subject area specializations, elementary/secondary school supervision, etc. Sometimes these charts told us how people performing these functions consulted with, or reported to one another, and to principals, teachers. Sometimes not. From the formal organizational arrangements, it was not possible to infer what districts were doing about instructional improvement.

The above partial listing of findings foreshadows the conditions that we now regard as essential, even if not sufficient, to the creation of a management strategy to link testing with instruction.

Three conditions that were present in varying degrees and manifestations in all six of our districts, and which we infer to be sine qua nons of a data-based instructional change management strategy were:

1) Motivation--that is, the presence of some strong impetus or collection of incentives;
2) Idea champions--that is, the presence of leaders or a critical mass of others who have knowledge and interest in both tests and instructional improvement, and who occupy positions from which to mandate or persuade others of the legitimacy and likely payoff of this approach;
3) Delivery system--that is, coordination among competently run operations within the central office; and communication channels between the central office and the schools.

We'll take these essential conditions one at a time and describe what we found in our six districts, four of which had been evolving their strategies for more than eight years, two of which were relative newcomers with only a two year history. Needless to say, most districts regarded their efforts as "in progress," even those that appeared to us most successful.

1. Motivation: impetus and incentives. In each district respondents told us that they as professionals were concerned and frustrated by the low levels of student achievement as evidenced by test scores. Their frustration was often echoed by parents, the

media, and the school board. However, the specific catalysts which turned this general concern into district activities differed from district to district. They included combinations of the following:

a. Court directives to raise test scores as an indication of district good faith in providing equal educational opportunity;

b. Federal and state evaluation or testing requirements accompanying categorical program funding;

c. State requirements to develop competency testing programs;

d. Availability of federal or state grants to develop basic skills tests;

e. Board policies directing the establishment of fundamental schools district-wide;

f. Board policies directing the development of criterion-referenced testing systems;

g. Parent pressure for higher rates of admission into college, better vocational preparation;

h. Influential district staff--committed by previous graduate training or in-service professional contacts--to a test-teach-retest-reteach instructional cycle.

2. Idea champions (Daft & Becker, 1978). We cannot overemphasize the importance of what we have come to call, familiarly, the "care-clout" factor. In each of our districts there was someone, either by him or herself or with a small group of colleagues, who cared--and persisted over a long time in that caring-- about using either the tests themselves, or the student's scores on the tests as a level to improve instruction. This person was not necessarily the highest official in the district. In one district, a relatively

low status administrator allied himself with the Board president to promote his ideas (and ultimately himself). In another district, the Director of the research and evaluation unit who pushed this test-instruction linkage later became the superintendent. in a third district, a new superintendent allied himself with a very well-liked supervisor of curriculum to provide the direction and energy for initiating a district-wide strategy to link testing with instruction. But, in every case, the idea champion had or soon acquired formal or informal clout.

In our six districts, the idea champions did not seem to plan or implement their activities in a goal directed fashion. Rather, they had a general vision towards which they were driving, encouraging others to make use of opportunities as they occurred. They did one or more of the following:

a. Legitimized and shaped the informal as well as policy decisions concerning data-based instructional improvement;

b. Found allies among opinion leaders within the central office and teacher and principal groups;

c. Reinterpreted or reconceptualized the district's past and present activities so as to provide the rationale for future activities. This process of reordering some of what had already been done was especially noticable in districts' writing of proposals for new funds;

d. Mobilized energy, raised morale, and transformed feelings of staff and teacher helplessness into feelings of empowerment;

e. Restructured the rewards and sanctions within the district.

3. Delivery system. In many districts, operations of units which which carry set functions relating to curriculum, instruction, supervision, administration, personnel, budgeting often operate

autonomously. Staff members rarely have formal or informal opportunities to share their problems, perceptions of goals, etc. with one another. By the presence of a delivery system in our six districts, we mean not only the performance in competent and timely fashion of activities connecting testing with instruction. We also mean the existence of coordinating mechanisms--whether they be meetings, memos, informal conversations--that insured the meshing of activities at the district, school and classroom levels. Our six districts ranged widely in the number of divisions, units, schools or classrooms which were either centrally or peripherally involved in data-based instructional change. They also ranged widely in the formality and frequency of intra-organizational arrangements for talking about, deciding on, implementing or monitoring data-based instructional change.

Here is an example of one district which seemed to us to have a well-developed delivery system. Here, the individuals--often a combination of staff, principals and teachers--responsible for developing the curricular scope and sequence, for constructing the criterion-referenced tests, for organizing staff development courses, for ordering new books and media, for hiring new teachers, for developing budgets, interacted frequently with one another on an informal level, talking with one another in the halls and in one anothers' offices. They also had frequent and regularly scheduled meetings. Each respondent reported regarding the promotion of student learning on specified objectives through the test-teach-retest-reteach cycle as a strong influence on the way

> in which he or she carried out job related duties.
>
> In another larger district, by contrast, the delivery system was less well developed. For example, the development of criterion-referenced tests was handled within one unit, the norm-referenced testing program was handled by a second unit, and staff development was carried out independently of these activities. Interaction among individuals even those with shared concerns, was largely accidental, depending sometimes on previously formed friendships or associations. Although contact between central office staff and principals was frequent and formally scheduled, several principals told us that they often heard conflicting stories from different central office unit personnel.

In our six districts, then, we found that there had been specific incentives to raise student achievement levels by using the tests, that there were idea champions intentionally moving these ideas into action, and there was some form of delivery system wherein the activities undertaken by various units of the district were coordinated with one another and communicated to relevant audiences.

Management Strategies

Our districts were engaged in many tasks, each of which might be regarded as an isolated activity to connect testing with instruction. Rather than list the activities individually, however, we have grouped them into configurations. Each configuration or strategy represents a more or less coherent management orientation within districts. Some districts had "clean" management strategies--using only one orientation--while others had "mixed" strategies, that is, pursuing simultaneously activities which seemed to belong to several configurations. The three strategies are:

1) A personnel-improvement-oriented staff development strategy;
2) A building-oriented problem solving strategy;
3) An instructionally-oriented objectives-based strategy.

1. A personnel-improvement-oriented staff development strategy. One of our districts used this strategy explicity in conjunction with the objectives-based strategy. Another advocated this strategy as the key to data-based instructional change. In this latter district, central office staff reasoned that the student population within each of their schools was heterogeneous and becoming more so, that teaching and principal staffs were stable and likely to remain so, that teachers themselves made the major difference in student learning, so

- District officials wrote grants for federal and state money to conduct district-wide inservice.
- They integrated the construction of state proficiency tests with staff development courses, training teachers in writing objectives and items.
- District staff checked district-constructed state proficiency tests as they were developed against high school course offerings, found skills which were not being taught, organized teacher committees to develop materials, provided staff development to teachers newly assigned to teach that content.
- District staff analyzed students' State Assessment Program test scores by subscale, checked textbooks against subscale content checked teachers' instruction time against subscale content and organized staff development courses for particular teachers on how to teach those identified skills.
- District staff required all teachers' attendance at courses on how to teach using diagnostic/prescriptive techniques, where the diagnosis was to be informed by students' test scores.
- District staff required principals' attendance at courses on supervision and on diagnostic/prescriptive instruction; and then

mandated that principals were to spend 50 percent of their week in classrooms observing and facilitating instruction.

2. A building-oriented problem solving strategy. Two large districts were using this strategy. Their central office staff reasoned that the schools in their districts, by reasons of history, geography, or present ethnic populations, represented distinctive organizational entities. The principal, teachers, parents, students and surrounding community were regarded by themselves and by the district as the primary actors responsible for improving students' learning; therefore, these individuals together should be identifying problems and devising solutions, assisted by whatever district support seemed advisable. So (and these examples come primarily from one of the two districts)

- The evaluation branch provided to principals, teachers and Title I coordinators printouts of norm-referenced (mandatory) test scores as well as results of the annual School Information Survey.
- The Curriculum and Instructional branch distributed to schools the criterion-referenced (voluntary) test scored.
- The evaluation office appointed local school evaluators whose responsibility was to interpret to principals and their staffs the results of norm-referenced test data. These local school evaluators sat in, where requested, on beginning of the year school planning meetings where school level goals for the year were made based on areas of need identified from test score patterns. During the year, local school evaluators responded to principal, teacher and parent advisory board requests for test interpretations and instructional directions to pursue.
- The evaluation office encouraged in a pilot set of volunteer schools a process called local school budgeting which involved parents, teachers and principals in data collection and analysis activities designed to inform the school's allocations of its annual budget revenues.
- Area supervisors asserted (although without monitoring or sanctions) the principals' responsibility for using these printouts in school site planning and in conference with individual teachers about classroom management and about individual students.

3. An instructionally-oriented, objectives-based strategy. Two districts, one large and one small, seemed to be using this strategy. Each had started approximately eight years ago. Each came to adopt a highly structured diagnostic/prescriptive instructional model supported by a scope and sequence outline of objectives, a criterion-referenced testing system under continuous revision to keep it updated and de-bugged, media and materials cross-referenced to objectives and to the tests. In one district, not only was there a tight connection among curriculum, instruction and testing, but there was also

- Compulsory staff development for principals, teachers, aides volunteers and substitutes during school hours. Between sessions teacher-taught model lessons within the classroom were observed by the staff development coordinator and the principal.
- Released time for teachers and principals to attend conferences on instruction and teaching.
- Weekly district-wide principal meetings to discuss individual school and across district problems.
- Clear delineation of roles and responsibilities from board members through to aides, with follow-up and monitoring of performance of one level by the next higher level.
- Support resources for teachers in the form of a learning specialist available to help plan classroom management based on CRT printouts, work with individual children.

In neither of these districts did the objectives-based orientation imply top-down decision making. Rather, in both, there was a high level of communication and involvement between operations in the central office, as well as a high level of participation of teachers and principals in thinking about, doing and reflecting on data-based instructional change.

It should be noted that, conceptually, each of these strategies could have been managed in either a tightly or a loosely coupled manner (Weick, 1976), thus making six possible strategies. The differences between tight and loose coupling would show up most clearly in the feedback and monitoring aspects of the coordinating mechanisms which exist among central office operations and between them and the schools. The two districts which used the instructionally oriented management strategy appeared to us to be more tightly coupled than the other four, but this may not be an inevitable accompaniment of a particular strategy.

Characteristics of Management Strategies

What we have said so far is this. In our six districts--where there has been a publicly acknowledged intention to move in the direction of data-based instructional change--there also has been some relatively specific impetus or incentive that stimulated the process; one or more individuals who have acted as idea champions; and some set of district structures which coordinated their individual action in relation to linking testing with instruction. District-wide strategies to link testing with instruction seem to be oriented in one of three directions: towards staff development where the emphasis is on influencing individuals' attitudes and behaviors; towards local school buildings where the emphasis is on involving school staffs in data-based problem solving; towards administratively-oriented tight coupling where the emphasis is on a minimum set of clearly-defined instructional objectives.

We'd like to offer some impressionistic characterizations of these strategies.

Uniqueness. Although we ourselves found it possible to generalize about configurations or strategies, we were struck during our visits by the uniqueness of what each of our six districts was doing and how they explained their reasons for their activities. That is, the handling of linking activities in each district seemed to have been influenced by idiosyncratic factors such as local history, local geography, the image of the district in the eyes of the public and of the people who worked there, local politics both within the community and within the district, immediate events, crises or funding availability. Especially important seemed to be the personalities of and power relationships among the people within the district. Although we tried to avoid it, we could not help observing to one another the cliches about "education being a people business" and "people matter." It seemed to explain much of the variability among districts.

Non-exclusiveness. A second characteristic that occurred to us was the non-exclusive nature of what districts were doing. The strategies for linking testing with instruction, while important in the minds of many of most of our respondents, was only one of the involvements and concerns that occupied their workday; and sometimes other crises or problems sidetracked, either for a few days or for much longer periods the concern with data-based instructional change.

Additionally, no district had what might be called a blueprint or a masterplan for this particular subset of concerns. Some individuals, in two of the districts, expressed their sense of what the data-based instructional change jigsaw puzzle might look like once all the pieces were in place. We found the jigsaw puzzle metaphor to be a useful one. In some districts, we could infer thet most of the boundary edge pieces were identified along with many of the inside pieces. In other districts, there may have been large pieces on the table, but there seemed to be available, as yet, no straight-edged boundary pieces to enclose them.

Episodic. Finally, and clearly related to the preceding point about the piecemeal nature of the strategies, is our observation about the evolution of the strategy itself. Instead of being linear and sequential--that is, instead of proceeding in an orderly way from planning or organizing, implementing, evaluating and recycling--the management of data-based instructional change was episodic and moved on a broken front. Activities speeded up or slowed strategies down in accordance with deadlines or other scheduled events. Implementation--that is, action--often took place in the absence of any explicitly stated plan. Formal plans were sometimes generated after the fact in order to explain the actions that had occurred.

Implications

We will be spending the next year working with district representatives on a Guidebook for managing data-based instructional change. In advance of this work, we would not want to elaborate all the implications of these observations for school districts who want to do something about data-based instructional change, but we can make some obvious points.

1. It appears that district-wide management of data-based instructional change can and does occur. Some districts have moved a long distance towards management strategies in which testing is linked to instruction in ways that are intended to improve student learning.
2. It appears that any data-based instructional improvement change process is complex and slow to evolve. It requires people

with skill and knowledge not only in the substantive aspects of testing and instruction but in the management aspects of conceptualizing, organizing, directing and monitoring.

3. It appears that local factors and local people are critically important in shaping the strategies which districts use to manage data-based instructional change. Although there are generic issues and cross-cutting conceptualizations which can be identified by research and by experience and which would be helpful for district personnel to know, there is likely no simple standardized formula which districts can follow. Instead, districts, having decided that this is a course they want to pursue, must get all the help they can assemble and then build their strategy out of locally-available ingredients.

References

Daft, R. L., & Becker, S. W. The innovative operation: Innovation adoption in school organizations. New York: Elsevier, 1978.

Guba, E. G. Toward a methodology of naturalistic inquiry in educational evaluation. Los Angeles: Center for the Study of Evaluation, University of California, 1978. P. 6.

Kennedy, M. M., Apling, R., & Neumann, W. F. The role of evaluation and test information in public schools. Cambridge, Mass.: The Huron Institute, 1980.

Weick, K. E. Educational organizations as loosely coupled systems. Administrative Science Quarterly, 1976, 21, 1-19.

Assessing the Effects of District Testing and Evaluation Efforts

We have been working with the problem of how to capture the effects--intended and unintended, direct and indirect--of a district's activities linking testing and evaluation with instruction. Our interviews with teachers, based on our understanding of district practices proved frustratingly uninterpretable. We then realized we might have to adapt our impact studies to a whole range of district intentions some of which were articulated, some not; and that we would have to discriminate effects based on levels. We experimented with the "stakeholder" notions of Mason and Mitroff, but were prevented from further study by unexpected events in two case study districts.

Clearly, this paper is a first effort at delineating the problem. Likely, we will develop a series of short papers, some theoretical and some more practical, which will expand the ideas suggested here. With luck, they will prove provocative enough to include the forthcoming Management Guidebook.

Assessing the Effects of District Testing and Evaluation Efforts

Donna Mitroff

Introduction

Over the past two years we have intensely studied six school districts which were identified as involved in activities to link their testing and/or evaluation activities with instruction (T, E and I). We documented what they were doing, how they came to be doing what they are doing, how they were set up operationally and how they thought all of these efforts linked testing with instruction. From the outset of the study, we anticipated that stage of the research which asks, in effect, "So what!"

The "so what" question deals with the impact of the testing and evaluation activities of a given district. In the original project proposal, we expressed the issue as that of understanding "the impact or effect of district-wide testing and evaluation activities on the actions of teachers and principals in classrooms and schools." (CSE Plan, 1979, p.18.) We wanted to assess the extent to which the TEI linkage subsystem was having the "desired" or "expected" effects in classrooms.

We have spent considerable time during the past year examining the topic of T/E impact assessment. Through dialogue, review of related work, and some pilot applications we have reformulated the original issue, refined our definitions of T/E impact and outlined a procedure for others to use for themselves in clarifying their thinking about T/E impact. The purpose of this paper is to share our progress in working through these fundamental methodological issues.

Addressing the Issue

Our original statement of the issue, by focusing on "desired or expected effects in classrooms," suggested that we would look directly in classrooms for teacher behaviors which would indicate that they are selecting and performing actions based on input from testing and/or evaluation data. The first step we should take in designing our research, then, would be to spell out what we would look for as evidence of such effects. In order to do this, we examined what districts were doing in testing and evaluation and attempted to develop reasonable scenarios for what the common impact might be. We asked ourselves: "Given this district's particular testing and evaluation subsystem, what types of effects would flow from it and be evident in classrooms?"

The first thing we learned was that we were confused by the terms we had been taking as synonymous. "Effects," "Impact," and "Use" are not terms that we could continue to use interchangeably unless we wanted to stay hopelessly confused. To clarify our purposes, we adopted the definitions suggested by Smith (1981) for the terms "use" and "impact." Smith defined "use" as "conscious employment of an evaluation (or test) to achieve some desired end or impact," and "impact" as "any discernible actions, events, or changes in conditions that are directly influenced by the evaluation (or testing activities), its processes, products, or findings." Extending these definitions, we add that uses are intended effects whereas impacts can be either intended or unintended. That is, examples of either "uses' or "impacts" can both be referred to as "effects" of a T/E/I linking subsystem.

Returning to the effort to specify indicators, we built a matrix as in Table 1.

TABLE 1: Specifying Effects

	Examples of			Sample Indicators
	Uses	Impacts		
Tests		Intended	Unintended	(a) All GATE candidates have 98% CTBS scores
1 - CTBS 2 3 . n	(a) Scores will be used to identify candidates for the GATE Program	(c) Teachers identify potential candidates for GATE, notify special services.	(b) Teachers use results to assign reading groups	(b) Examination of classroom reading group profiles shows homogeneous grouping by CTBS scores. (c) Check in cum. cards indicates GATE children who were cited by teachers.
Evaluations				
1 2 3 n				

Using the data from our field work in six districts to try out this matrix led us to the delineation of several key dimensions concerning the assessment of TE Use/Impact. They are listed here in random order.

- Testing and evaluation use and impact can both occur at many levels of the district's organization in addition to the classroom.
- Testing and evaluation use and impact may be understood differently at different levels of the organization, e.g., administrators focus on test scores in relation to an entire program while teachers see scores in relation to individual children.
- Among the variety of observable impacts, some may be intended by someone in a decision-making position; some may be unintended.
- Many of the likely T/E effects are not recognized or articulated by members of the organization.
- There may be a lack of consensus among members of the organization over what constitutes "acceptable evidence" that effects are in fact occurring.
- Among the varieties of effects which can be described, some can be categorized as a directly technical, that is, instructional, emphasis while others reflect a social/interpersonal emphasis which may or may not indirectly effect instruction.
- Among the varieties of effects some are experienced inside the organization while others are felt in the environment outside the organization, e.g., in the community, by the media, etc.
- Neither anticipated use nor anticipated impact are typically built into the organization as clearly as one would expect.

Assessing the intraorganizational and the environmental effects of testing and evaluation in instruction requires a much broader perspective than that which we had first anticipated. Not only must effects be sought in places other than the classroom; not only must effects on instruction include use as well as intended and unintended impacts; not only must the direct effect on instruction be accounted for, but also the indirect effects mediated through the social/interpersonal processes.

We must therefore step back to the level of school district ideas, policies and practices. We need to find a way to assess the effects of both the articulated policies and practices of the district related to the use of testing and evaluation for instructional change as well as the unarticulated intentions. Much of the confusion contributing to multi-level assessments of the effects of testing and evaluations stems from the fact that many district attitudes related to testing and evaluation are not stated as a consistent policy position, but are evolutionary, reactive to circumstances (e.g., public outcry at declining scores), or dependent on the preferences of those in key opinion leader positions.

Therefore, defining the effects that testing and evaluation are supposed to have is not a simple matter of asking one or more policy makers or of searching for a written statement of district policy. Ideas, policies, practices, and expectations change as they filter through the organization and through people's perceptions. The effects--that is, the uses and the impacts of testing and evaluation--occur differentially at different levels as this filtering process occurs.

District intentions regarding testing and evaluation, then, are dynamic in the sense that those who serve as transmitters of intentions are also adding or modifying the original intentions. This modification occurs at all levels of the organization and suggests that the transmitters or agents of intended policy are also at all levels of the system. To put it another way, ideas and policies are defined and implemented by all those who have a stake in them...all such individuals can be called policy "stakeholders."

This stakeholder concept is of central importance not only to those of us who would like to assess the effects of testing and evaluation on instruction but also to those within a district who would like to manage a testing/evaluation/instruction linkage.

Testing and evaluation are activities which should be carried out with reference to (and deference to) stakeholders. Stakeholders are defined as "those claimants inside and outside of the system who have a vested interest in the problem under investigation and its solution" (Mitroff & Mason, 1981). Those working with the stakeholder concept ask: Who is affected; who has an interest; who is in a position to effect adoption of results or execution of decisions; who has expressed opinions; who ought to care about outcomes? The stakeholder concept is related to the previous statement that effects occur at many levels and are defined differently at different levels, and further complicated by the fact that effects occur both inside and outside of the system. Stakeholders, therefore, may be either internal or external to the school district organization.

To summarize our thinking thus far: the strands which come together are these...

- one cannot effectively assess the effects of a district's testing and evaluation activities without an understanding of the goals and intentions of the district;
- district goals and intentions are embedded in the ideas and in the policies and practices of the district;
- some of the ideas and policies are implicit rather than explicit--unarticulated, rather than articulated; practices may be either consistent or inconsistent with the prevailing ideas and policies;
- the effort to make the ideas and policies explicit must involve a wide range of stakeholders at all levels of the organization.

A Process for Use/Impact Clarification

Over the past several months, our project staff has been developing and pilot testing a process to elicit a school district's intentions related to the effect of testing and evaluation on instruction. This process uses a structured workshop in which participants from many levels of the school district organization collaborate.

In addition to enabling a school district to surface the implicit expectations of stakeholders at many levels within the organization as well as those outside the organization, it is our hope that the process can be a planning device for districts seeking to create a T/E/I linking system.

The workshop format has two principal justifications. First, it allows participation of individuals from many levels who have diverse perspectives, reflecting our belief that knowledge resides at many levels of the system. It is therefore not sufficient to explore use/impact intentions either for planning or for assessment with input from only the members of a testing and evaluation unit. Second, workshops can incorporate procedures which build "ownership" in ideas or policies. The dynamic participative workshop procedure that we are constructing, hopefully meets these conditions.

Our workshop requires a minimum of 8, a maximum of 16 people from across levels and functions in the district. It calls for at least 2½ hours and is even more comfortable if conducted in a longer session.

The workshop procesures are adapted from Mason and Mitroff (1981) who have applied their methodology for dealing with "ill-structured problems"* in both public and private agencies. We have incorporated their procedures into five major steps which are first listed then discussed in detail.

1. Generate examples of effects of testing and evaluation.
2. Determine the importance and certainty of effects of testing and evaluation.
3. Specify acceptable evidence of effects.
4. Select those effects to be measured.
5. Develop instrumentation to measure effects.

1. Generate Examples of Effects

The basic procedure used in this step is brainstorming -- a somewhat structured form of brainstorming known as the nominal group technique. The nominal group technique requires that the group facilitator go from one person to the next in turn, asking each person to contribute one or more ideas to a group list which is being compiled. The use of nominal group process simply insures that each person in the group has an opportunity to contribute at least one item to the list. The process of going nominally from one participant to the next is continued for successive rounds until members of the group begin to pass. When the point is reached that no other items are forthcoming the process is terminated.

The group facilitator begins the process by asking for examples of some effects of testing and evaluation which participants have observed. No effort is made to focus or channel the items at this point.

---

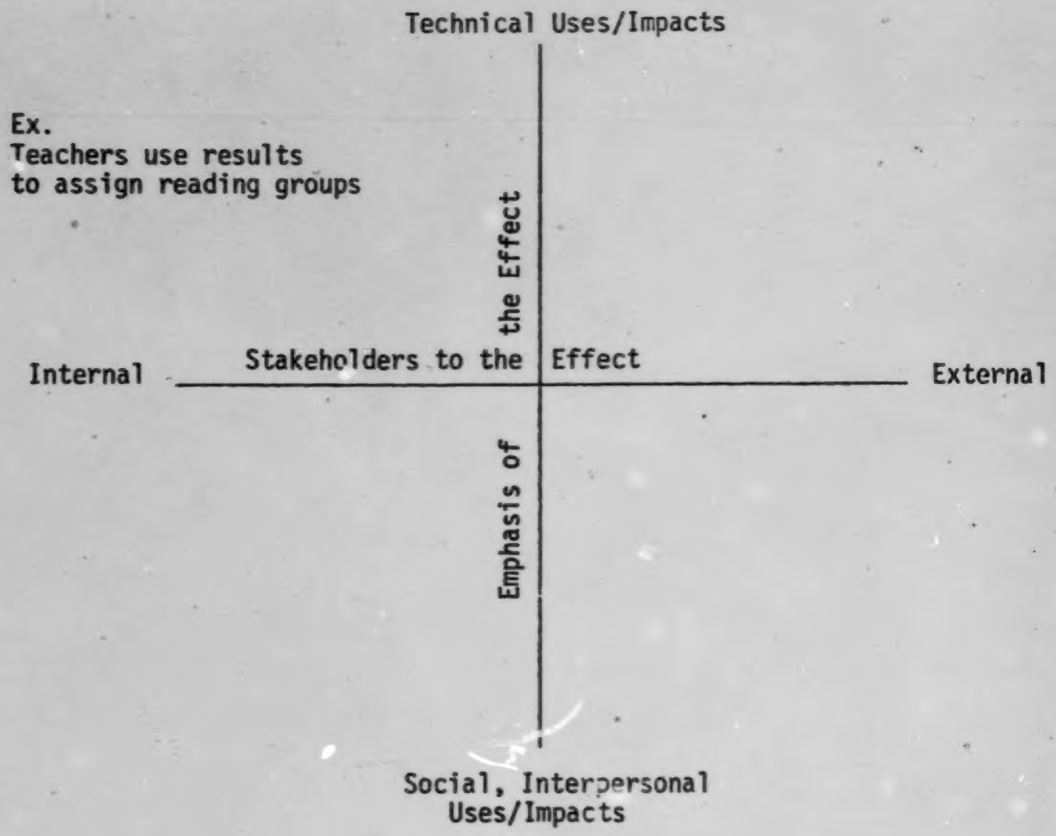
*Ill-structured problems are defined as those for which "there are no single right answers; there is no consensus even on the definition of the problems; and action steps will or should be taken in spite of these ambiguities" (Mason & Mitroff, 1981, p. 29).

Furthermore, no effort is made to judge items contributed to the list.

When it is clear that the brainstorming process has reached an end, that is when items are no longer coming forth from the group, the facilitator moves on to a sub-step to clarify what has happened in the group and possibly stimulate more thought by clustering the items. A useful clustering strategy can be illustrated by the matrix in Figure 1. We have found that examples generated by these brainstorming sessions can be roughly classified according to two dimensions. The first dimension to be considered is the stakeholders to whom the examples apply and whether those stakeholders are internal or external to the school. The other dimension along which we find a great number of examples clustering is the emphasis of the effect. The emphasis of the effects tend to be either technical or social/interpersonal.

When we take these two dimensions and overlay them to form a matrix, the resulting four quadrants combine different stakeholders with different effects. As an example of the use of the classification scheme, consider one of the examples from Table 1, "Teachers use results to assign reading groups." The stakeholders (classroom teachers and their students) are internal and the effect is technical, i.e., a technical type of instructional decision. This effect has been placed in the upper left quadrant.

Figure 1. Framework for Clustering T/E Effects


Technical Uses/Impacts
Ex.
Teachers use results
to assign reading groups
Internal
Stakeholders to the Effect
External
Emphasis of the Effect
Social, Interpersonal
Uses/Impacts

In pilot applications of the process we have drawn this figure on a blackboard or flip-chart and illustrated to workshop participants how their examples can be grouped in one or the other of the quadrants. It is not necessary to classify every item on the diagram -- a representative sampling is sufficient. The facilitator then asks the group if the framework suggests other examples of effects which could be added to the list. Oftentimes it will, and when those are added, the facilitator senses the time to move into step 2.

2. Specify Acceptable Evidence of Effects

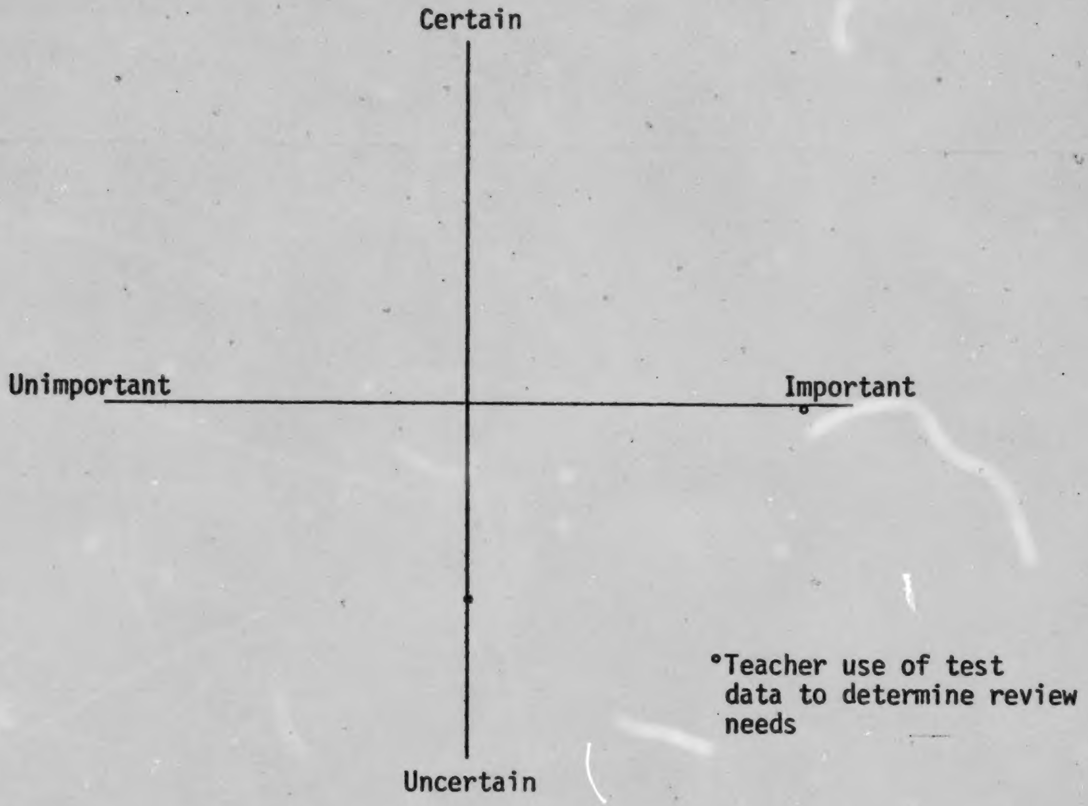
In order to carry out this step the group is broken up into sub-groups. Sub-grouping is helpful because it provides an opportunity for more interaction and more contribution from group members. It also permits participants to consider a selected sub-set of effects rather than the entire range.

To subdivide the group, participants are asked to select themselves, in fairly equal numbers, into one of the four quadrants. One group will then be considering the internal-technical effects, another group will concentrate on the external technical, a third group will consider the internal social-inter personal, a fourth will consider the external social interpersonal.

Once these sub-groups have been formed, the assignment for each group is to take all effects from the brainstorming phase which they feel justified in placing in their quadrant and consider the importance/unimportance, certainty/uncertainty of each of those effects in terms of the district's overall testing and evaluation effort. Once again, participants will be using a cross matrix, as illustrated in Figure 2. This time the horizontal line is scaled from important to unimportant while the vertical line is scaled from certain to uncertain. Each effect is discussed and placed on the classification scheme according to the consensus of the group.

By way of example, consider a sample effect such as the following: test results used by remedial reading teacher to determine consonant blends to be reviewed. The group may agree that such use or impact of test results is important and place it far out on the Important dimension. They may

Figure 2. Determining Importance Certainty


Certain
Unimportant
Important
Uncertain
°Teacher use of test data to determine review needs

disagree, however, on the certainty of the effect. One person thinks it is a very isolated occurrence; another saying that many teachers use the results in that way. The group decides to place the effect in the important/uncertain quadrant. In so doing they have had "flagged" it as a potential topic for further study and have helped to clarify what it means to them.

Once the sub-groups have completed this step they have, in essence, completed a first sorting of effects and made explicit their notions about which testing and evaluation effects are both "important and uncertain." Effects which fall into that quadrant of the classification scheme are regarded as the most critical effects according to the judgment of the particular small group. Each sub-group presents their classifications to the entire group. At this point there may be discussion and some reorganizing of priorities. More important though, each effect has been systematically considered.

3. Specifying acceptable evidence of effects.

Once again, participants work in sub-groups. This time each of the subgroups is instructed to take those effects which they labeled as "important and uncertain" and produce for each an example of evidence which might be useful in reducing the uncertainty about that particular effect.

For the example used above, the group may decide that interview data indicating that 2/3 of the remedial reading teachers used test results to plan their review activities would be acceptable evidence.

When each of the small groups have completed this task, a general group session is convened for each sub-group to present their list of effects and corresponding examples of evidence.

The purpose of this step is to involve school district personnel in the specification of data sets which they, themselves, will find acceptable. The step is designed to help prevent us, as researchers, from designing and conducting a research study which can be summarily dismissed by its intended clients.

4. Select Effects to be Measured

The effects and evidence lists from each group are compiled into one complete list. It is likely that the composite list is too long for the time and resources of most research. One way to pare it down is to have workshop participants rank order the total list for research priority. The ranking can be used to determine which effects should become the subject of continued research.

5. Develop instrumentation to measure effects.

This step requires that the statements of evidence be used to develop instruments to collect data to prove the presence or absence, strength or weakness of selected effects. In our case the instrumentation is designed by our group, the research team, and presented to the school district team for reaction and revision. Again, the intent is to involve the clients of the research in its design.

The effects clarification process ends at this point. However, the collaborative climate of the process needs to continue through the implementation of the research.

Discussion of Pilot Applications

We have conducted two trials of the clarification procedure. The first was a simulation using members of our Center staff in school district roles. This trial was devised to enable us to try out, revise, and refine the agenda. The second was in one of the sample districts, Northtown, wherein we intended to proceed on to develop instruments and assess T/E effects. Recent events in the school district, however, precluded that opportunity and we proceeded only the point of developing a set of research recommendations.

A third trial in another of the sample districts was planned but again events in the district (relevant environments) were such that it could not be carried out.

Staff simulation:

Five staff members assumed the roles of Director of a Testing and Evaluation Unit, District Superintendent, principal, elementary teacher, and secondary teacher. All of the persons assuming roles had, at some time in their careers, worked for a public school system and were familiar with school issues.

In order to set a context for the simulation, participants read a sample district case study and assumed that scenario for their roles.

The trial, abbreviated by the fact that we had only 1 1/2 hours in which to conduct it, was carried out in our conference room. The experience suggested revisions to the process. These revisions are reflected in the preceding discussion and will not be elaborated here. Instead, we present the content outcomes of the process.

Table 2. Summary and Classification of Effects from Simulation Pilot

Stakeholders to effects

Emphasis of effect	Internal	External
Technical uses/impacts	° Helps teachers make better use of statistics ° Builds curriculum rigidity ° Enables teachers to speak a more common language ° Promotes contact with parents for home management ° Kids are grouped ° Meets individual instructional needs ° Takes up teaching time ° Counselors use to program students ° Provides needs assessment information	° Provides achievement test data for Board ° Accountability function ° Enables school to receive funds ° Provides reports to parents
Social-interpersonal uses/impacts	° Kids get grouped for socio-metric purposes ° Some kids get upset by tests ° Promotes contact with parents on non-instructional as well as instructional information ° Focuses on cognitive learning ignores affective domain ° Intensifies competition between teachers and between schools	° Newspaper reports to community ° Gets greater/lesser pubkic support ° Consultation with parents ° Brings recognition from outside sources

Table 3: Listing of Important/Uncertain Effects and Examples of Evidence from Simulation Pilot

Important/Uncertain Effects	Evidence
• Kids are grouped for socio-metric purposes	• Examine individual teacher grouping patterns; interview teachers for rationale.
• Focuses attention on cognitive over affective domain	• Observe classroom lessons; examine lesson plans; examine teacher questioning patterns.
• Builds curriculum rigidity	• Interview teachers; examine teacher's objectives over time.
• Enables schools to receive outside funds	• Interview Assistant Superintendent; interview funding sources about why they funded district.
• Brings recognition to district from outside sources	• Interview educational colleagues, applicants for positions in district. Interview researchers at university re district's reputation.

All participants rank ordered the effects listed in Table 3. Their rankings indicated that, were this an actual case rather than a simulation, the research would address the following issues as priorities.

1. Builds curriculum rigidity: (Does the testing and evaluation activity have this effect?)
2. Focuses attention on cognitive over affective domain: (Does the testing and evaluation activity have this effect?)

3. Kids are grouped for socio-metric purposes (Do teachers use T/E data in this way?)
4. Enables school to receive outside funds (Have funds become available because of T/E activity?)
5. Brings recognition from outside sources (Can it be shown that the T/E activity has positively increased the district's visibility?)

We point out that while numbers 1,2,3, are classroom based effects, numbers 4 and 5 are the types of effects which can have an indirect effect on classroom instruction.

Because this was a simulation we did not proceed to Step 5 - Develop instrumentation to measure effects. However, the results of the process prepare the research team for that step by providing a client centered focus.

Comments: Participants in the simulation, all experienced with general educational issues and with specific T/E issues, felt that the process brought out aspects of testing and evaluation activities which they had not considered. One of the participants described the process "a series of sieves through which the issues get refined and focused." They felt the interaction was particularly helpful to their new understandings.

Pilot Application in Northtown.

Through the Assistant Director of the RD and E Unit a two hour session was arranged in Northtown. Discussions with the Assistant Director prior to the workshop determined that the concentration during the workshop would be on the consolidated application process. That process, defined in another project report (ED Project Annual Report, Nov 1980 p 82) can be described as follows:

- Each consolidated-application school's CTBS scores obtained initially and presented to each school's principal and staff along with the school's mobility index, minority percent, and school enrollment figures;
- Based on these data, the school staff, with the assistance of an Evaluation Services Office evaluator, determines a set of objectives and activities for the coming year. These form the core of the school's annual improvement plan. District evaluators regularly revisit these schools during the ensuing year. The CTBS tests are administered again in the Spring and individual pupil results are reported to the appropriate teacher before the end of the school year. During the Summer, the Evaluation Services Office staff scores the tests and analyzes the results in terms of the individual school's stated goals. A school-specific report is prepared and presented to the school staff in the beginning of the Fall quarter. This forms the basis for the school staff to reformulate goals and activities for the next year -- and the cycle is repeated.

Eight school district representatives participated in the session: there were three from the RD & E unit, two from Title I programs, one principal, and two resource teachers.

We opened the session with a brief summary of the results of our research and an indication of the next phase -- assessing effects. It was clear from the start of the session, and in fact it had been anticipated by the Assistant Director prior to the session, that there were many different agendas on the minds of participants. We also knew that the district was in a state of anticipation of a potential court ruling on desegregation. The many unresolved issues and emotions absorbed some of the allotted time and the full process was not completed.

Table 4: Summary and clustering of Effects of the Consolidated Application Process in Northtown

Stakeholders to the Effects

Emphasis of the Effects

	Internal	External
Technical uses/impacts	• Report comes out too late to affect activities • CRT's are used by teachers • Low CTBS scores led to strict classroom interruption rules • Process has no bearing on what teachers do in classrooms	• NRT results used by SEA • RD & E staff conducts in-service when principal cannot. Gets the RD & E staff into some schools • CTBS analysis led to in-service and special materials • It satisfies reporting requirements
Social-interpersonal uses/ impacts	• Certain minority groups score poorly because of language problems • Principals set up in-service for PR reasons • Testing takes up too much student time • Children are burned out from over testing	• Parents get a better view of what is happening because of report • Parents in many schools don't understand the report • Public relations from report is good for some schools - bad for others • Newspaper publishes NRT results • Public is fixated on CTBS scores

Beyond the generating of the effects, the process got bogged down and time had run out. When we reflected on what the session told us about the effects of the consolidated application process, we concluded that:

- the process is having its impact mainly in the external technical and external/social-interpersonal area, but very little impact on the internal areas. It would be wrong to even consider effects on classroom activity because such effects have not been built into the consolidated application process;
- the level and form of participation in the consolidated application process is different from school to school and any effort to assess the process needs to use the school site as the unit of analysis;
- the principal's behavior and attitude will be key to level of use in a given school site.

These thoughts along with our suggestions for data collection procedures were shared with the district personnel.

References

CSE Three Year Plan proposal to the National Institute of Education. Los Angeles: Center for the Study of Evaluation, University of California, 8/30/79.

CSE Evaluation design project: School district organization Study submitted to the National Institute of Education. Los Angeles: Center for the Study of Evaluation, University of California, December, 1980.

Mason, R. O., & Mitroff, I. I. Challenging strategic planning assumptions. New York: John Wiley, 1981.

Smith, N. L. The president's corner. Evaluation News, Summer, 1980, No. 16, p. 24-25.

Changing Teacher Behavior: From Symbolism to Reality

Richard C. Williams

This paper was prepared for a Working Conference on Changing Teacher Practice, Austin, Texas, October 15-16, 1981. Publication is expected through the Conference sponsors. A journal for elementary school administrators and teachers has also requested permission to print.

This paper draws on the concepts presented in CSE Monograph # 10, Evaluation in School Districts: Organizational Perspectives and the data gathered in the project's six districts. It argues that the decline in public satisfaction with schools coupled with the emergence of research connecting teaching with learning in a more predictable manner than heretofore calls for an organizational rather than an individual approach to instructional change.

Introduction

There seems to be little question that American public education is presently facing a serious crisis. The origins of the crisis are many; some are societal, e.g., declining population, rising social disorder with the resultant crime and vandalism, shifting priorities that divert funds away from social programs. These and other external conditions and developments have decreased the resources available to education and complicated the already difficult tasks that the public schools have been called upon to perform.

But another, internal, cause of the educational crisis is the public schools' malfunctioning. The popular press, e.g., Time (1980), reports numerous instances of teacher incompetency and administrative inability to efficiently and effectively deliver educational services, and there has been a concomitant decline in student achievement.

At this working conference we are examining one facet of the internal problems facing the schools; that is how to improve teacher performance and the role inservice training might play in such improvement efforts. To be sure, ineffective teacher classroom behavior is but one part of the problem. Other factors, such as inept administrators and student antisocial attitudes, contribute in turn to teacher ineffectiveness. Because of the influence these other factors have, it seems unlikely to assume that teacher classroom performance can be or will be improved by teachers alone. Thus in my paper I will attack the problem as not only a teacher's responsibility but also a school district's responsibility.

In this paper I want to present a perspective, a point of view, rather than a fully developed argument. There are two reasons for my tentativeness: one, my thinking about how best to improve teacher behavior has been changing over the last year or two and I have not yet fully formulated these new thoughts, and two, I understand the spirit of this conference to be one of inquiry and exploration--it would be inappropriate to present and fiercely defend a particular viewpoint. You represent various roles in the educational system--I look forward to your reactions.

Let me state a few caveats and warnings. I will often refer in general to American Education. As you know, that enterprise--American Education--is enormously varied, largely decentralized, and very complex. Generalizing about American Education is a risky business at best. I realize that there are likely many exceptions to my generalizations. Also, my paper is based on assumptions underlying work in progress on teaching effectiveness. My interpretation of this work may be limited--I welcome your comments. Finally, my argument may challenge assumptions and beliefs held by some of you. I am not trying purposefully to be provocative; my purpose is to try to shake all of us out of orthodox thinking--something I think necessary if we are going to make progress in improving teacher behavior.

Let me briefly outline the paper that follows. First I will describe why school districts can be described as institutionalized and, therefore attend to teacher behavior and teacher effectiveness in essentially a symbolic way. Then I will discuss some recent developments that may allow school districts to deal with teacher behavior more effectively and thereby become more like a technical system, and finally I will discuss the implications this new development might have for school district management.

School Districts as Institutionalized Organizations

A common observation of American public school teachers is they work quite autonomously behind closed classroom doors (Lortie, 1975). To be sure teachers are occasionally visited by principals and other supervisors, especially during their probationary term, but for the most part teachers are largely unsupervised when it comes to the day-to-day interactions they have with their pupils. Similarily, the relationships between each teacher's methods and his/her pupil's learning, are largely unknown or can at best only be estimated (Goodlad, 1974). Thus we have a curious situation where what would seem to be the most important interaction in a schooling system, that is teacher classroom behavior and its effects on pupils learning, is largely unattended to in any direct supervisory manner in most school districts. This is even more surprising when one views the reportedly low level of preparation teachers generally receive before they begin their difficult task (Lortie, 1975). (In California presently there is a bill before the Legislature to require teachers to pass a written test on basic reading and mathematics skills.)

The now fashionable concept of organizational analysis called "loose coupling" (Weick, 1976) helps describe this phenomenon. "Loose coupling" means that in organizations, "units, processes, actions and individuals are typically connected loosely rather than tightly. Thus actions by one agent or element have little predictable relationship to the actions of another element or agent" (Clark, 1980). Most school districts, when it comes to managing their most important interactions--namely the interaction between teachers and pupils and the effects thereof--are loosely coupled (Meyer, 1977). One manifestation of this loose coupling is that teacher behavior becomes self-determined and both effective and ineffective teachers continue their work largely unmonitored by anyone from outside the classroom door. It doesn't seem likely that continuing this pattern will result in any significant improvement in teacher performance.

Why are teacher interaction patterns with pupils and their results only loosely coupled from organizational management? A number of explanations have been offered for this phenomenon. One of the most persistent views expressed whenever school administrators gather at local, state, and national meetings is: "I really want to be an instructional leader but all the paper work and other problems distract me from this important task" (Williams, Hill, & Wuchitech, 1980). Accordingly, principals take time-management classes and make solemn promises to mend their ways--but teachers remain isolated behind their classroom doors. Others (Goodlad, 1978) feel that school administrators have been going through a phase (a second era) in which they have allowed purely administrative, and therefore peripheral, matters to divert them from their central task, which, in Goodlad's view, should be the school's curriculum and instructional program. Goodlad joins the growing chorus of voices calling upon administrators to reestablish instruction as a top job priority.

In a somewhat similar vein, Erickson (1975) feels that educational administration research and training programs have gone awry, focusing as they do on organizational theory, collective bargaining, politics, economics, and the like. They give far too little attention to important questions about what instructional methods and supporting organizational structures will result in the greatest pupil learning. Like Goodlad, he urges a "paradigm shift" in which educational administration refocuses its attention and priorities toward instruction.

Given these rather persistent statements about the need for administrators to attend to instruction, one would expect that today we would find principals and principal preparation programs turning their attention in this direction. I suspect, however, that the "paradigm shift" Erickson calls for has not taken place except in a symbolic manner (Dornbusch & Scott, 1975). School administrators largely do not attend specifically to day-to-day classroom instructional activities. If this is so, what explains this reluctance of school districts and school administrators to address this important concern?

In my view, one of the most penetrating analyses of this phenomenon has been expressed by Meyer and Rowan (1977). They confirm the observation that school districts, when it comes to their technical core, that is the careful monitoring of teacher instructional methods and its result on pupil learning, are indeed loosely coupled. But, they note, it is a mistake to label the typical school district's instructional supervision as totally loosely coupled. Indeed, school districts have many instruction-related activities and standards that are very tightly coupled. They point out that districts pay very close attention to such things as: whether or not teachers are credentialed; whether schools meet various accreditation-related standards such as providing the appropriate number of library volumes per pupil, limiting classroom size, having a published curriculum, and providing inservice training for teachers and administrators. Whether the teachers indeed teach the established curriculum or whether or not credentialed teachers are using sound instructional methods, however, receive little direct district attention other than routine teacher or administrator workshops or inservice training.

Meyer and Rowan regard these activities and similar accreditation-related standards as essentially symbolic acts which, when added together, constitute what the public accepts as an adequate definition of education and what school administrators call instructional improvement. Very little is done by school districts to assure that changes actually take place in classrooms or that teachers achieve the desired results. School administrators, when asked whether or not their schools are good or not, will often answer in terms of symbols, such as the percentage of teachers with advanced degrees, quality of facilities, and soundness of the curriculum, and the decision-making processes. Meyer and Rowan label organizations that behave in this way as institutionalized, as opposed to technical organizations. Institutionalized organizations tend to tightly couple their symbols to their organizational structure; conversely technical organizations tightly couple their organizational structure to their technical core.

Meyer and Rowan (1977) offer several reasons why school districts have adopted this institutionalized mode. I will discuss three which are relevant to the perspective I am presenting.

1. Education has a remarkably weak technical core. They assert educators really know little about the relationship between teacher behavior and pupil learning--unlike hospitals or many manufacturing firms for example, which have a relatively strong technical core. It is very difficult and even dysfunctional to tightly link an organizational structure to a weak technical core.

2. School districts over the decades can be thought of as having been very successful, with no need to attend to the technical core. How can school districts be considered successful? If providing a real and equal opportunity for all children to meet their potential is the criterion--then schools seem clearly and increasingly unsuccessful. However, if the success criteria include steadily increasing budgets, government protection, and a fairly high general public opinion of schools, then schools indeed have been quite successful. For example, school districts over the years have shown consistent budget growth and the curriculum has increased from the basic core to embrace such subjects as driver education, nutrition, and sex education. The number of people employed in school systems has grown steadily over the decades. Clearly, the public has supported the public education with all its institutional characteristics. In terms of organizational survival, the American public school clearly has been a winner!

3. Another reason for the institutionalized character of schools, they argue, is that its behavior is consistent with its assigned role as a social and economic sorting machine. The fact that some children do better than others is, in the public mind, to be expected because not all children are equally endowed with the abilities to meet the demands of schooling--and society. Some are winners, some are losers; and the school cannot be blamed for this. There is no need to look at the technical system because it has only limited power to influence pupil achievement.

Thus in Meyer and Rowan's view the institutional nature of schools is an enlightened response by sensible people to their situation. If schools are successful as they are; why muck about with an unsure technology and perhaps display the weaknesses of the technical core? In the public's mind, a school is a school is a school. So, in spite of their protestations to the contrary, school administrators are not inclined, or professionally able, to pay close attention to their technical instructional core.

In many ways, I am impressed with the Meyer and Rowan thesis. In light of recent developments in education, however, I think there is reason to question whether or not this institutionalized mode must persist. In my opinion, school districts can and should become more technical organizations. Let me clarify that observation by critiquing the Meyer and Rowan analysis; I have reservations about their third reason, and conditions in schools may have changed so as to allow districts to change from their institutionalized posture and become more technical. First, I will express some reservations about their third reason, then make comments on the first two. Meyer and Rowan are not alone in arguing that school districts respond to society's need for a "sorting machine" and are therefore not really interested in looking closely at the interaction between teacher behavior and pupil achievement. Some economists (Bowles & Gintis, 1976) have maintained that the schools have helped sort out people into various economic classes. Speaking of the development of a supply of skilled labor, they state:

> Indeed we shall suggest that the maintenance of such a "reserve army" of skilled labor has been a major, and not unintended, effect of U.S. Education through the years.

Personally I have some difficulty with this as an explanation for the lack of school district attention to the technical instructional core. One reason for my doubt is the complexity of the American educational system. The argument of those who see some purposeful national policy to deliberately sort people into social classes has to rest on the assumption that there is some group who can effectively influence an entire diversified national schooling system. I have no doubt that some unknown, and perhaps large, percentage of school districts have fulfilled that sorting role. No doubt many continue to do so today. But I believe that many educators do wish to provide an equal opportunity for all their students to achieve at their highest level and they are frustrated by their inability to do so.

I don't think lack of school district desire to improve pupil achievement and life and career chances fully explains the institutionalized character. What may have influenced the problem has been a lack of a technical core--school administrators simply have not known what teacher behaviors and instructional methods most affect and improve pupil achievement. Lacking that core, their behavior has become institutional, they have tried to maintain public support through symbolic acts. And until quite recently that strategy has been quite successful. The efficacy of their symbolic acts had resulted in a generally high level of public support. The public had been willing to accept what the schools called education as satisfactory. School districts have not had to attend to their technical core.

But there have been two important changes in recent years. One change is a decline in the public's satisfaction with and support of the public schools. There are many reasons for this, e.g., perceived poor discipline, disagreement over desegregation solutions such as busing, changing opinions regarding whether or not the public schools should have a continuing semi-monopoly over educating America's youth. But a major reason, I would suggest, is the increasingly visible signs that the schools simply aren't doing a very good job, as evidenced by declining test scores in general and a persistent inability to assist those in the lower socioeconomic strata to improve their academic performance. In spite of all the schools' symbolic acts, the public is crying for better results. Sensing the continuing inability of the public schools to deliver, they are withdrawing their support.

The other important change I perceive is the emergence of a better understanding of the relationship between teacher behavior and pupil achievement which holds promise of developing into a technical instructional core.

The elements of that emerging technology are well known to many of you, and will likely be dealt with at some length by others here, so I will only mention them briefly. A major component has been the findings of the BTES research (Fisher et al., 1978) which identified a link between allocated Academic Learning Time and pupil achievement. I would add to that research, the work of those who have been further probing the effect of teacher behavior on pupil achievement and deriving the practical implications of that work (Stallings, 1980). Related to that is the work of those who have been, for several years, drawing upon principles of learning and classroom-proven methods to develop both the science and art of teaching (Hunter, 1971). I would also include the developing technology in competency-based testing which can provide a more direct measure between instruction and pupil achievement, and thereby serve as a valuable instructional tool (O'Shea, 1981).

This shift from instructional symbolism to instructional technology can be aptly illustrated by Bloom's description of mastery learning in which he describes the following implications mastery learning has for instructional practice: Available time versus time-on task (amount of time invested differs from the active learning time spent); intelligence versus cognitive entry (predicting academic achievement on the basis of intelligence measures differs from using specific knowledge, abilities or skills as prerequisites for learning); summative versus formative testing (using tests to judge a student's grade versus using tests as a diagnostic and instructional tool); teachers versus teaching (judging teacher quality on the basis of traits and training versus judging teacher effectiveness by examining instructional behaviors); parent status versus home environment conditions (estimating pupil success by noting home and family racial, socio-economic and cultural characteristics versus observing parent and pupil interactions and behavior in the home). (Bloom, 1981, pp 3-17.)

To be sure, a fully developed technology of teaching does not yet exist. Both Bloom (1980) and Stallings (1980) remind us that much additional research and developmental work remains to be completed. Moreover, Denham (1980) and Fenstermacher (1980) warn us about moving too quickly from the BTES findings to a overly definitive and rule-bound instructional system for all teachers. These are important caveats and caution is well advised, but I would disagree with those who would insist that the most that can be made of these findings is to develop materials for teacher workshops or for teacher use as their felt needs, discretion, or interests lead them to it.

Given the two significant changes in the schools' systems that were described above--the decline in the public's satisfaction with and support of the public schools, and the emergence of a technical instructional core--serious efforts to improve teacher performance calls for, I believe, an organizational response by the total school system. If there can be general agreement that some teacher behaviors and methods will more than likely result in greater pupil growth than will other teacher behaviors and methods, when school districts must begin to identify those behaviors and attend closely to whether or not they are being used by teachers. School districts must identify instruction and pupil achievement as a major technical responsibility. In other words, school districts should begin the task of tightly coupling their organizational structure to an agreed-upon technical instructional core. This means that they must abandon the approach of offering "cafeteria style" inservice training and instead develop ways to assure that the agreed-upon instructional practices are utilized. They must re-examine the attitude that improving teacher performance is soley the responsibility of the classroom teacher. The entire school district, teachers and administrators, should begin exploring ways to link their organizational structure to the instructional core.

To continue to ignore this crucial activity will, in my opinion, result in a further decline in pupil achievement and a continuing erosion of the public's confidence in the schools.

School Districts as Technical Organizations

How would a school district as a technical organization differ from school districts that are institutionalized organizations? Essentially it would mean that it would no longer leave specific teacher classroom behavior and the measurement of its effects solely to teacher discretion, behind the classroom door. The school district would likely provide inservice training on desired teacher behaviors, and they would see to it that teachers actually behaved in the desired manner, and they would determine whether or not the desired behavior was having its intended effect.

Let me illustrate this more specifically. A colleague of mine, Adrianne Bank, and I are conducting a three-year research study at UCLA's Center for the Study of Evaluation on how school districts can link testing and evaluation with instructional improvement (Williams & Bank, 1981). Based on our observations and analyses, we believe that some components of a technically based school district instructional program would have certain components. I will not attempt to present some blueprint that would be common for all school districts; we have learned that there are many ways in which school districts differ on conditions or variables that can critically influence the design and effective implementation of such a plan. No common program will work everywhere.

An important, perhaps the most critical, component is a comprehensive idea of what the district wishes to accomplish. What, for instance, is the overall logic of the plan? What parts of an instructional renewal plan does the district want to emphasize? Are there common teacher behaviors or conditions the district wants to see occurring in each classroom? What district conditions, specifically and generally related to schooling, impede or help the district? For example, it makes a difference in terms of inservice training needs if the district is already staffed with a large percent of competent, experienced teachers rather than a large cadre of new, inexperienced teachers. Additionally, teachers traditionally work alone behind closed doors; they do not readily embrace new teaching technologies into their repertoires. How will you bridge this barrier in attempting to change deeply embedded teacher behavior?

A district must determine its unit of change; by this I mean, where it focuses its attention--at the school building, the district level, or some combination thereof. In the past many have argued that the most effective change strategies should be built around the school site, with the principal playing a key change-agent role. In some instances, this would be most appropriate. In spite of the widespread support for this strategy, however, I have increasing reservations about its use as the only way to change and improve teacher behavior. I believe our faith in principals' ability to change teacher behavior is sometimes overestimated. One reason for my reservation is that experienced teachers generally have little faith in the ability of a non-teaching principal, who often has less training and classroom experience than some teachers, to offer valid advice. Also, there simply hasn't been a technical core with agreed-upon techniques and vocabulary to which principals could turn when trying to change teacher behavior.

With regard to differing strategies, some school districts may agree at the district-wide level about an agreed-upon set of preferred teacher techniques and behaviors. In such districts the principal's task may be to assure that these standards and behaviors are being exhibited by teachers. But note that this would shift the principal's role from that of an expert who can determine teacher effectiveness to that of being an agent who assures the implementation of district-wide standards that have been agreed upon by teachers and administrators. In other districts, it may be more appropriate to consider the school as a unit, with the principal playing a major instructional role.

Another critical factor would seem to be the district's determination to enforce its instructional renewal plan. That is, some mechanism must be developed to assure that the district's agreed-upon standards are indeed being implemented. Specific steps must be taken to tightly couple the organizational structure to the technical core. Someone has to supervise the supervisor!

A consequence of this is that districts will eventually have to take appropriate action to assure adequate compliance or provide incentives to encourage those who must meet district standards and to assist those who cannot. An admittedly difficult problem is determining appropriate steps to be taken with those who cannot or will not show adequate improvement. Typically district teacher contracts and tradition have all but elminated the termination of contracts of incompetent teachers. And to be sure it will likely be a long time, if ever, before a sufficiently developed technical core will be available that could be used for determining teacher competency. If the link between teacher behavior and student achievement can be firmly established, however, it seems reasonable that all parties to the educational enterprise--administrators, unions, parents and the courts--will want to seek a fair and reasonable way to assure that such behaviors are encouraged.

A final component would be that the districts would develop a set of criterion-referenced tests that would be available to the teachers for diagnostic teaching purposes. These tests should be in a highly usable form, e.g., easily administered, quickly scored, and directly tied to the district's instructional program. The link between test results and their instructional implications should be developed so that teachers can make use of the results rather than have the results used against them (O'Shea, 1981).

There are numerous other components and conditions that likely must be attended to. Let me dwell on just one more, and a crucial one, namely the scope of responsibility. I do not view the development, and implementation of a technical core to be solely a top-down, management responsibility; it is a school district responsibility. School administrators are no more qualified nor motivated to improve instruction than are teachers or teacher unions. This effort must be a district-wide responsibility, with all the parties participating in the program's planning and implementation. Likely management, with its wider organizational perspective and responsibilities will initiate such a plan. But such a plan will not go very far if it is essentially top-down and ultimately develops into an adversary relationship between management and labor, that exludes the public. Developing and utilizing an effective technical core is everyone's business and in everyone's interest.

I have not talked directly about inservice training, but it is obviously a key to this approach because improvement will have to take place largely with experienced teachers. Inservice training would have to be geared to the special needs of each district and would include such activities as developing the technical core, teaching appropriate skills, developing and implementing a criterion-referenced testing program, and determining the implications of criterion-test results for instructional practice. A key factor is that inservice training be linked to a district plan rather than a set of management-determined, symbolic activities which may have only a marginal relationship to what will most improve teacher behavior and pupil learning.

Is It Possible?

Perhaps some of you are thinking, "Well, these are interesting theoretical notions, but it is quite unrealistic to think that this can ever be put into practice." To this I would reply that two of the six districts that Adrianne Bank and I are studying, have been for years developing and implementing plans like those described. (Also see Stow, 1979.) One is a large urban district, the other a small suburban district. While following somewhat different paths to get to their present positions, both have defined a technical core, provided for the widespread understanding of the teacher behaviors that are expected, provided necessary inservice training for teachers and supervisors, provided for direct supervision (tight coupling), and related their efforts to competency-based pupil achievement measures. Far more than most districts, they have the qualities of a technical system--their technical core is tightly coupled with the organizational structure. It has taken approximately eight years for the districts to reach this point. There is still much work to be done; however, the districts already are convinced that this approach has resulted in increased pupil achievement.

Let me conclude with some speculations about ways in which this shift the from an institutionalized system to a technical system might occur in districts. Several things might cause it to happen. One, some "idea champion" or cohort of reformers, may recognize the opportunity and push in this direction simply as a way of improving the educational system. In one of the districts we studied, this was the way it happened. The other way it may come about is when districts are so persistently unsuccessful in improving pupil achievement that they begin to lose public support and their very existence is threatened. They may turn to such a strategy as a crisis solution.

If a teaching technical core is emerging and its use in classrooms is determined to have an important impact on pupil achievement, then we cannot wait for a full-blown technology to develop before we act. We should seriously consider taking what is available and still developing and fashioning a technical core, however limited, in school districts, and implementing it in a fair and humane way. I maintain this will only happen effectively when changing teacher behavior is acted upon as a school district responsibility.

References

1. Bloom, B. All Our Children Leaving, New York: McGraw Hill, 1980.
2. Bloom, B. "New Direction in Educational Research and Measurement Attenable Variables," Invited address, Annual Conference of American Educational Research Association, April 15, 1981.
3. Bowles, S and Gintis, H. Schooling in Capitalist America: Educational Reform and the Contradiction of Economic Life, New York: Basic Books, 1976.

3a. Clark, D. L. In consideration of goal-free planning: The failure of traditional planning systems in education. In D. L. Clark, S. McKinnin, & M. Malkas (eds.), New perspectives on planning in educational organizations. San Francisco: Far West Laboratory, 1980.

4. Cohen, E. G., Deal, T. E., Meyer, J. W. and Scott, W. R. "Organization and Insturction in Elementary Schools," Technical Report No. 50. Stanford: Stanford Center for Research and Development in Teaching, 1976.
5. Denham, C. H. "Lessons for Policy Makers," Ch. 14 in Time to Learn, National Institute of Education, 1980. Pp. 231-238.
6. Denham, C. H., and Lieberman, A. (Eds.), Time to Learn, National Institute of Education, 1980.
7. Dornbusch, S., and Scott, W.R. Evaluation and the Exercise of Authority, San Francisco: Jossey-Bose, 1975.
8. Erickson, D. A. "An Overdue Paradigm Shift in Educational Administration, or Thow Can We Get that Idiot Off the Freeway," Paper presented at Educational Administration Twenty Years Later 1954-1974," Ohio State University, April 27-30, 1975.
9. Fenstemacher, G. D. "On Learning to Teach Effectively From Research to Teacher Effectiveness," Time to Learn, National Institute of Education, 1980. Ch. 14, pp. 231-238.

10. Fisher, C. W., Berliner, D., et al. "Teacher Behaviors, Academic Learning Time, and Student Achievement: An Overview, Ch. 1 in Time to Learn, National Institute of Education, 1980, pp. 7-32.

11. Fisher, C. W., Filby, N. N., Manilave, R. S., Cahan, L. S., Dishaw, M. M., Moore, J. E., and Benliner, D. Teaching Behaviors, Academic Learning Time and Student Achievement: Final Report of Phase III-B, Beginning teacher Evaluation Study, Far West Regional Laboratory, San Francisco, CA 1978.

12. Goodlad, J. I. "Educational Leadership: Toward the Third Era," Educational Leadership, January 1978, pp. 322-331.

13. Goodlad, J. I., Klein, M. F., and Associates. Behind the Classroom Door, Worthington, OH, Charles A. Jones, 1974.

14. Hunter, M. "The Teaching Process and the Learning Process," in Handbook For Teachers, Seifman E. and Alken, D. (Eds.). Glenview: Scott Foresman, 1971.

15. Kirst, M. W. "Curriclum: A Key to Improving Academic Standards," Paper presented at the College Board Symposium on Transition from Secondary School to College, St. Louis, MO, May, 1981.

16. Lortie, D. Schoolteacher, Chicago: University of Chicago Press, 1975.

17. March, J. G., and Olsen, J. P. Ambiguity and Choice in Organizations, Bergen: Universitetets Forlaget, 1976.

18. Meyer, J. W. "Research on School and District Organizations," Paper presented at Sociology of Education Conference, San Diego, CA, April 1977.

19. Meyer, J. W., and Rowan, B. "Notes on the Structure of Educational Organizations," Chapter in M. Meyer, et al., Studies in Environment and Organization, San Francisco, Jossey-Bose, 1977.

20. O'Shea, D. "School District Evaluation Units: Problems and Possibilities," Monograph No. 10, UCLA center for the Study of Evaluation, Los Angeles, CA. Adrianne Bank and Richard C. Williams (Eds.), 1981. (In press.)

22. Stallings, J. "Allocated Academic Learning Time Revisited, or Beyond Time on Task," Educatonal Researcher, December 1980, pp. 11-16.

23. Stow, S. B. "Using Effectiveness Research in Teacher Evaluation," Educational Leadership, Oct. 1979, pp. 55-58.

24. Weick, K. E. "Educational Organizations as Loosely Coupled Systems," Administrative Science Quarterly, pp. 1-19, March 1976.

25. Williams, R. C., Bank, A. "Uses of Data to Improve Instruction in Local School Districts: Problems and Possibilities," Chapter in Applying Evaluation to Education, Carol Aslanian (Ed.) Beverly Hills, CA., Sage Publications (Forthcoming 1981).

26. Williams, R. C., Hill, P., and Wuchitech, J. The Effects of Federal Education Programs on School Principals, Rand Note, Prepared for U.S. Department of Health, Education and Welfare, February, 1980.

APPENDIX

CRESCENT CITY

1981

CRESCENT CITY

1981

BACKGROUND INFORMATION

The Setting

Crescent City is a large school district (enrollment over 80,000 pupils) and as such, it shares many characteristics with other comparably-sized districts. For example, it has a growing minority population (Black, 15 percent; Hispanic, 5 percent; Asian, 2 percent, and American Indian, 1 percent). These percentages are smaller than are found in many American urban centers but the district is experiencing a steady growth in its minority population. The district has implemented a court-ordered desegregation plan.

The district is facing an increasingly tight financial situation. The most recent state legislature session almost passed a proposition 13-type bill to limit property tax increases. Even without that, the state has increasingly stifled the sources of school district money from the local community to the state. In the 1960's the state's share of the budget was 40 percent--now it is 60 percent. The state presently has a fiscally conservative governor and legislature. All school board members and administrators with whom we talked were pessimistic about the ability of the district to balance its budgets in the near future without severe cuts.

It is difficult to see where these cuts will be made. The district ranks near the bottom nationally in terms of class size (large classes) and in expenditure per pupil (low expenditure). Crescent City has a higher cost of living index than the average urban city, and teacher and administrative salaries continue to slip behind the inflation rate. Faced with a projected low legislative budget increase for the coming year, the district is bracing for a potential teacher strike.

While Crescent City shares several characteristics with its urban counterparts, i.e., growing racial minorities, declining financial resources, large classes and low per-pupil expenditures, and growing teacher unrest, it has several unique characteristics. Its pupil population has shown a steady growth pattern since the 1960's. (It has added 17,000 pupils since 1970.) One possible nearby development may mean a city population growth of 50,000 people within a 2-3 year period of time. Thus, when most other urban districts are auctioneering off old school buildings, Crescent City is building several new schools per year and hiring large numbers of teachers whom it has recruited nation-wide.

Unlike other urban districts, the district has no nearby suburban school district that can drain off pupils or attract teachers for various reasons, e.g., White flight, better working conditions. Thus, it has a virtual monopoly over public education. There are several private and parochial schools; however, one of the city's major religious groups, while maintaining its own after-school educational program does not operate its own school system.

One of the city's major industries and the supporting business have considerable employee transiency. Many families move regularly in and out of town, and move repeatedly among the school district's various attendance areas.

School District Organization

Like many of its urban counterparts, the Crescent City school district has shifted periodically between centralization and decentralization of its administrative functions. In the 1950's the district was quite centralized. Beginning in the early 1960's, however, a charismatic superintendent led a rapid move towards decentralization. The district was divided into four subdistricts and individual school principals were encouraged to "climb their school's flagpole, assess the attendance area's special needs and develop a school program especially tailored to each school's pupil needs."

The result was a proliferation of programs and approaches and wide variations in quality among the district's schools. Beginning in 1970, successive superintendents, who were faced with community unrest over the schools' uneven quality and declining achievement levels, began to recentralize the subdistricts. The four administrative areas were eliminated and the superintendent, a powerful deputy, and four associate superintendents assumed great district-wide decision-making powers.

Since his appointment in 1978, the present superintendent has flattened the organizational chart further by eliminating the deputy superintendent position and appointing six associate superintendents who report directly to him. The associate superintendents' titles are: Personnel Services, Business and Finance, School Facilities, Elementary Instruction, Secondary Instruction, and Administration and Special Student Services (which includes the Department of Research and Development).

While at first glance this might appear very similar to other school district administrative arrangements, it should be noted that there is no separate department of curriculum or instruction that independently services the entire district. Instead the curriculum department has been folded into the divisions administered by the associate superintendents for elementary instruction and secondary instruction. The curriculum specialists and supervisors report to the top line administrators who, in turn, administer the elementary and secondary schools. The importance of this arrangement, which was initiated by the present superintendent, will be pointed out later in this case study.

Another relevant administrative-structural component are the Directors, who report directly to the Associate Superintendents for Elementary and Secondary Instruction. Each Director is responsible for a set of geographically determined schools. They are the administrative and supervisory extensions of the Associate Superintendent and they play a critical role in the District's instructional management program.

In addition to the Superintendent's cabinet, which consists of the superintendent and associate superintendent, there is an infrastructure of committees, including a principals' advisory committee and various curriculum advisory committee.

The district operates its own administrative training program which identifies and prepares district teachers and administrators for promotions to or through the administrative ranks. Few administrators are appointed from outside the district.

CURRICULUM AND INSTRUCTION PROGRAM

History

As was indicated previously the district has gone from a decentralized, diversified organizational structure and curriculum and instructional program to its more centralized present structure and program.

In our respondents' opinion, a major reason for this shift was because of considerable public concern over the school pupils' achievement on standardized tests. (In the late 1960's the percentile scores on the Metropolitan Achievement Test were in the high 30's and low 40's.) Around 1970, school board and public concerns were highlighted in the newspaper and the district began to take a close look at its decentralized and diversified curriculum. As a leading reformer stated, "We had a lousy approach, not focusing on the right things." Influenced by several nationally known advocates of management by objectives, competency-based instructional methods, and clinical supervision, some central administrators began developing a comprehensive instructional program that would improve the district pupils' achievement. It should be noted that this program was chiefly advocated by one person who was allowed to develop this competency and objectives-based system. The superintendent in the early 1970's was not particularly interested in instructional management; neither he nor the board advocated a particular approach. Instead, one "idea champion" (that is, someone who believed in a particular approach and persistently pursued its implementation), who was not at the cabinet level, worked with selected colleagues to develop the school district's instructional system. It should be further noted that the instructional system was initially designed for use in elementary schools.

Thus, the present instructional management program has evolved slowly through the efforts of one leader and a small cadre of supporters who jointly developed, tested and ultimately installed their instructional management system. It was not a carefully planned and developed districtwide program. And, those who initiated it worked quite independently of the superintendent and board, though to be sure, they did not oppose the plan.

The district's instructional program has been first developed in the elementary schools--it is now being applied in the secondary schools. It should be noted that the program was directed initially at the district's reading and mathematics program. The essentials of the management system have been expanded to all subjects but the CRT's have not yet been developed beyond the math, reading, and language arts areas.

Description of the Instructional Management Plan

At the heart of the program is acceptance of the idea that there is a technology of teaching and that certain conditions or practices (e.g., time on task, appropriate ways to introduce a concept or skill) will result in better pupil achievement. As one respondent stated it: "There may be more than one right way to teach but there are also some clearly recognizable wrong ways." For example, the district believes time on task is a critical element in maximizing pupil achievement. Some teacher behaviors will more likely increase time on task than others.

The desirable conditions and practices have been distilled into what are known as Elements of Quality. Initially introduced in 1972, the Elements of Quality consist of eleven criteria against which a school and the instructional programs can be judged. The Elements rest on three assumptions and contain eleven applications. They are as follows:

Assumption 1. Goals and objectives need to be clearly written and communicated.

Application:
A. Statement of educational principles
B. Elements of Quality
C. Course of study and curriculum guides
D. Special priority objectives (HPO's)

Assumption 2. Means must be provided and used to assess the degree to which objectives are attained.

Application:
A. Testing program
B. Checklists of observable criteria
C. Opinion surveys
D. Management audits (internal and external)

Assumption 3. All assessment should culminate in program improvement decisions.

Application:
A. Implied action statements in assessment reports
B. Priority plans for improvement
C. Evaluation based on results

Supporting this program is a whole collection of films, videotapes, instructional packages, and in-service training activities which inform and guide all teachers who are subject to the Elements of Quality. The program revolves around a series of district-developed tools--e.g., assessing pupil progress, assigning pupils to instructional groups, altering instructional methods. Teachers are to be able to demonstrate to supervisors that they are indeed using these tests in the prescribed manner.

Teachers through in-service training programs and principal assistance are also expected to be acquainted with various instructional methods, and to be able to demonstrate that they can use them appropriately. For example, they should know what the district considers the proper way to introduce a concept or to develop pupil motivation. They must also show evidence that they are adhering to the district instructional continuum.

The uniqueness of this system is its attention to enforcing the use of these Elements of Quality. While teachers can teach beyond the district continuum (after covering the required material) and use various instructional approaches (if appropriate), they do not have the freedom to "do what they think is best" if it violates the spirit of the Elements. Often in many districts new instructional methods are introduced and explained and teachers are urged to use these methods. But teachers are left on their own, behind the classroom door, to implement the new program. Through painful experience, many districts have learned that changing teacher behavior is very difficult and many teachers, for various reasons, choose not to implement or practice the new program or method.

DISTRICT MANAGEMENT SYSTEM

Crescent City has implemented a management system to overcome teacher autonomy and resistance to their prescribed instructional program. The main agents for enforcing the district instructional program are school principals and the district's directors, described previously. Principals regularly receive extensive in-service training in clinical supervision; they are thoroughly informed on the district's instructional continuum and they are charged to oversee the implementation of the Elements of Quality in their schools. Each is expected to spend a minimum of 40 percent of his or her time in classrooms supervising teachers and assuring that the Elements of Quality are being adhered to. Examples of their activities would be: to observe a teacher introduce a concept and record whether or not the teacher did it properly and to provide help and advice if necessary; to ask teachers to justify, with diagnostic test results, the placement of pupils in skill level groups. The principals are, in turn, accountable to the directors who periodically visit their schools. Part of the Director's responsibility is to see to it that the principal is adhering to the Elements of Quality. The teachers are evaluated on their adherence to the dictates of the Elements of Quality; and so are the principals. The teachers, tenured and probationary, are reviewed by the principals and the principals are reviewed by the Directors.

Each year the principals are rated, on a confidential questionnaire, by the pupils, parents and teachers. These ratings, coupled with the Director's observation, form the basis for principal ratings. Teachers and principals who cannot perform to the Elements of Quality are provided extensive opportunities to become skilled. Teachers, for example, get multiple ratings and analyses of their teaching from several supervisors and in-service training opportunities are made available by the curriculum specialists in their division (elementary or secondary). If after several opportunities for improvement they cannot or will not meet the Elements' standards, they are subject to dismissal.

It is important to note that, while NRT's, and especially CRT's form part of this instructional management process, teachers are not to be judged on their pupils' performances on such tests. The district has adopted the process and insisted that the instructional staff and supervisors implement it. The responsibility for the results rests alternately with the central administration and school board. The chief architect of this system stated clearly that the system will fall apart if teachers perceive themselves as being judged on the basis of their classroom pupils' progress. He believes that once that happens, the teachers will, understandably, begin teaching to the tests and the system's intergrity will be compromised. But, although it is clearly not the district's intent, it could be inferred from several interviews with elementary school principals that students' test scores are being used to measure teacher effectiveness.

These Elements of Quality are being implemented differentially. The most complete use, including CRT's, has been implemented in the basic skill areas in the elementary schools and the Junior High Schools. The clinical supervision and adherence to the district's instructional continuum has been used quite widely in the elementary schools, somewhat less in the Junior Highs, because the "idea champion" of this approach has been associated with the elementary schools.

The Elements of Quality program is just now being implemented in the high schools, having been introduced three years ago. At this level the program embodies nine elements that the teachers are expected to follow. They are:

1. goals and objectives from course syllabus
2. criterion-referenced test scores
3. use of computer-assisted tests (CAT's)
4. class profile sheets for reading, math, language arts
5. individual student record profile cards for reading, math, language arts
6. correlated media, textbooks, and instructional materials
7. assistance from media clerk
8. use of curriculum support systems (e.g., English Support System Level I, Level II, Level III)
9. use of proficiency test scores

At present, these Elements are being used on a broken front. They are being used more in elementary schools than in secondary schools. Even in elementary schools, we were told, their use is variable because principal skill and commitment differs as does the intensity of the various directors' supervision. Secondary teachers are being introduced to this concept and the features of the system are being explained to them. The entire system, including CRT's and proficiency tests, is now limited only to remedial reading and math teachers.* The reason for this is that the district has not yet developed these materials for the large numbers of subjects and levels of instruction found in the high schools. The main Elements that are now being used are nos. 1 (goals and objectives from course syllabus) and 6 (correlated media, textbooks and materials), 7 (assistance from media clerk) and 8 (use of curriculum support system).

The reason for the reduced use of the Elements in secondary schools, in addition to its newness and lack of proficiency and competency tests, is also the greater autonomy that secondary teachers traditionally expect. They consider themselves subject-matter specialists--the imposition of outside advice and an institutional management system will not likely be readily embraced in secondary schools.

The Supervision Cycle

The use of the Elements of Quality can perhaps best be understood by reviewing the annual cycle of how it is used by one elementary division director. Basically, the director meets with each of his assigned principals in June for the end-of-the-year conference where they develop the next year's High Priority Objectives (HPO's). The director assists each principal to establish HPO's for himself and his school. The director also uses teacher questionnaire results to check on the principal's effectiveness in managing the Elements of Quality; Elements 1-5 focus on instructional objectives and Elements 6-10 (6-12 for secondary) focus on managerial objectives. In addition, the director uses parent questionnaire results to check on the school's effectiveness. These data are used internally, for the director's and principal's use only, and no normative data across the district is compiled. In September and October, the director begins his formal school visits and confirms the HPO's for each school, each principal, and each teacher in his division. The October through December months are spent in formal and informal monitoring of the instructional program. A mid-year assessment of everyone's progress is made in January and/or February. At this time the director conducts formal conferences and classroom observations with pre-submitted agendas and feedback procedures. For example, Form 440 is used to document recommendations made to each principal. March and April are spent in more formal and informal monitoring of the instructional program with data collection and verification. The in-service cycle for staff members assigned to the Special Assistance Program (those who received unsatisfactory evaluations) is completed. Around the end of April, the director compiles the data for his end-of-year report. His internal audit includes his own self-assessment, teacher-school profiles, assessment of instruction (Form 550), and his findings, conclusions, and implied action recommendations. The external audit compiles test results, opinionnaire results, division reports (audits), conference summaries, mid-year assessment, notes from school visitations, assessment of employee performance appraisals, and the 440 recommendations. In May he analyzes the data and completes his reports. During the end-of year evaluation, he shares his assessment with each principal. Together they relate this to the relevant HPO's, and establish tentative HPO's for the next school year.

---

*CAT's (Element # 3) are available in other subject areas as well.

Attitudes Toward the Instructional Program Management System

a) By the community:

We have no direct measures of community opinion. Our respondents, including school board members, implied that community criticism of the schools had abated now that the test scores had shown improvement. One might conclude that if the public related the test score gains to the new instructional management system, then presumably they would be supportive of the instructional system. At no time did anyone mention community concern that the system was overly prescriptive or that it unduly restrained teachers' instructional choice.

b) By the board:

The board members we talked with seemed impressed that the district administration had a definite plan for instructional management. They all were supportive of the superintendent's skill at and commitment to instructional improvement. The improved test scores have eased community pressure on the school board. To the extent that the board relates the test score gains to the instructional management system, it is likely that they feel partial toward it.

c) By principals and teachers:

A non-random sample of principals expressed mixed opinions of the program. Their opinions seemed to be influenced by many factors and it is difficult, without a more systematic measurement, to assess their opinions. For example, beginning teachers felt the program gave them much needed guidance, a critical aid not usually provided to teaching novices. They knew what to teach and what was expected of them. Some more experienced teachers felt the program's requirements and practices were not essentially different from what they had always done. Others, however, felt it was an imposition and required a lot of unnecessary and disruptive paperwork. The principals, especially at the elementary and junior high school level, seemed quite supportive. The program was so new and was being implemented in so few classrooms in high schools that the principals had few opinions to express. The teachers' association president was strongly opposed to it and felt that his opinion was shared by a large percent of the association members.

Perceived strengths and weaknesses of the Instruction Management System. One teacher and administrative respondents expressed the following opinions about the instructional program:

Strengths:

- Provides guidance for new teachers. Many teachers new to Crescent City acknowledged their increased security in the classroom provided by the IMS. Considering the past growth of Crescent City and its influx of new teachers annually, this is of considerable importance.
- Establishes criteria by which principals and teachers are judged. Principals and teachers know what they are being judged on.
- Provides basis for defining competence, identifying incompetence, and taking corrective action. Incompetent teachers

can be identified, offered in-service training and removed from the classroom if competence is not achieved.

Weaknesses:

- Increases paperwork for teachers and administrators.
- Catalyzes fears that teachers will be judged on students' achievement test results. The teachers' union advocates this position.
- Produces anxiety in teachers.

Linkages Among Testing, Evaluation, and Instruction

- Acceptance of a central idea ("There is a technology of instruction") links all parts of the instructional program.
- The technical core is tightly coupled with the organizational structure:
  1. Curriculum falls under the responsibility of a line supervisor (either the Associate Superintendent of Elementary Instruction or the Associate Superintendent of Secondary Instruction).
  2. Adherence to the Elements of Quality is maintain by Directors who report directly to this line supervisor thus providing a strong linkage between policy and practice.
  3. Communication system is functional since all involved parties meet on a regular basis.
- Supervision cycle provides strong linkage as evidenced by the "High Priority Objectives" selected by principals and teachers and monitored by Directors:

1. Criterion-referenced tests provide basis for instruction.
2. Norm-referenced tests provide basis for communication internally and externally.
3. Evaluation results from district-sponsored research studies provide data used to alter existing policies and programs. For example, a study on comparison of achievement test scores led to creation of the Elements of Quality; a study on student attendance patterns led to adoption of a district-wide attendance policy.

- In-service training is conducted by site principals for school staffs as well as by central office personnel from the appropriate division (elementary or secondary).
    1. Additional retraining is provided for staff members falling below expected performance levels.

Impact on Classroom Practice

From our visits to a non-random sample of elementary classrooms we noted several indices that the program had penetrated behind the classroom door. For example, teachers displayed on their classroom walls the test scores upon which they grouped students for reading instruction.

THE DISTRICT TESTING PROGRAM

The Research and Development Division administers the testing program. The district uses numerous tests in its testing program. They can be divided up several ways, e.g., mandated/non-mandated; subject matter/basic skills. For this case study's purposes they will be divided into criterion-referenced tests (CRT's) and norm-referenced tests (NRT's). Only those tests that are appropriate for understanding the testing, evaluation and instructional subsystem will be discussed here. Both norm-referenced and criterion-referenced tests are given. The following is a list of those tests:

Norm-Referenced Tests

- The Otis-Lennon School Ability Test

  (Given to all students in grades 2 & 5; been given in district for at least 6 years; originally selected "at least partly through R&E departmental evaluation of available products"; is given "for providing baseline data and for teacher information for individual student application.")

- The Stanford Achievement Test

  (Given to all students in grades 3 & 6; been given for four years; was selected to conform with the test given in the rest of the state; is given "as a performance indicator, statistics for minimal proficiency testing and to identify areas of strength and weakness.")

- The California Achievement Test

  (Given to all students in grades 8 & 11; been given for three years; was selected by a "Task Force of Teacher Consultants"; is given "as a performance indicator - teachers and principals are expected to use results to identify areas of strengths and weaknesses.")

- The Otis Lennon Mental Ability Test

  (Given to all students in grades 8 & 11; has been given in the district for 8-10 years; was originally selected "at least partly through a departmental evaluation of available products"; is given for "teacher information, individual student application and baseline data.")

NRT Scoring/Reporting. The district uses the "ACORN" statistical package to give the following information:

1. District and school comparative data:

   Data is given in the form of mean percentile scores for each school and the district for each subtest and for an overall score.

2. District and school stanine frequency distributions.
3. Percentile frequency distributions.
4. Raw score frequency distributions.
5. Statistical summaries (district by sub-test, sex, quartile).
6. Individual score list (on a gummed label).
7. Item analysis.

The district attempts to get results back to principals for distribution to teachers within 3-6 weeks. Comments from field people (i.e., people in the schools) indicate that it generally takes longer than this.

Actual uses of the Norm-Referenced tests. (1) A major use of NRT scores is to communicate to the community at large, to the Board, and to parents, how the district's students are doing. The community sees NRT scores when they are published in the newspapers. The Board, therefore, is also especially interested in these scores. Individual student scores are communicated to parents, usually in individual conferences where a counselor "interprets" the scores.

(2) A major second use of NRT scores is by the central office staff who use NRT information to examine the effects of district-wide instructional programs or policies. Low NRT scores were a major reason for the initial development and implementation of the current instructional management system. Additionally, the continued implementation and expansion of

the system is sanctioned by rising NRT scores. For example, the R&D Director said that "one reason we have moved the instructional management system to secondary schools are the blatent differences between elementary and secondary test results."

The scores reportedly began rising in the elementary grades, where the management system was implemented initially. As the management system has gradually been implemented in the junior and senior highs, those level's test scores have reportedly also begun to rise.

Some actual improvement of test scores on the NRT's can be documented. Longitudinal comparisons, however, are limited because the district administered the Metropolitan Achievement Test until 1977, at grades 2 and 5; thereafter they administered the Stanford Achievement Test at grades 3 and 6.

Scores on the tests were as follows:

		Metropolitan Mean Percentile Rank	
Grade	Year	Total Reading	Total Math
2	1974-75	70	62
2	1975-76	70	66
2	1976-77	72	76
5	1974-75	48	50
5	1975-76	52	54
5	1976-77	54	62
5	1977-78	54	70
5	1978-79	58	--

		Stanford Achievement Test Mean Percentile Rank	
Grade	Year	Total Reading	Total Math
3	1977-78	62	66
3	1978-79	64	68
3	1979-80	61	66
6	1977-78	54	56
6	1978-79	56	60
6	1979-80	60	60

As can be seen from the charts, not much can be said about changes for second and third grade reading scores. Second grade math scores increased dramatically on the Metropolitan, while third grade scores on the SAT didn't change much. Fifth grade Metropolitan math and reading scores both increased significantly, and sixth grade SAT scores went up in both math and reading.

We have no longitudinal data to support the district's belief that junior or senior high scores are now beginning to increase.

Another example of central office staff use of NRT information was an R&D unit study of the scores. The R&D unit director analyzed 8th and 11th grade NRT scores by quartiles and discovered that above average students are performing below their counterparts in the national norming sample. He speculates that this may be because of the "great efforts to address the state-mandated proficiency exam requirements," and the district's "failure to address the particular needs of above average students." Some of the central office directors are currently discussing the results of this study and the implications for modifying district instructional procedures.

(3) A third district use of NRT scores involved their utilization as part of an individual student's "index score." An index score is a composite of seven factors, including aptitude and achievement test scores, previous grade level, and other factors never delineated. Indexing is used primarily for placing students in certain levels (tracks), and for planning of a school's courses (for example, how many basic and advanced classes will be offered). At least one junior high principal, however,

uses the actual NRT scores (instead of the index score) for these placement and scheduling decisions.

(4) A fourth use of NRT scores, "taking corrective action" at the school level, was identified by both the director of high schools and the director of junior highs. The high school director claimed that her principals do an achievement versus ability study and an item analysis study of the NRT's to "determine what is lacking in their school programs." The junior high director said that they "zero in on areas of deficiencies that are also in the syllabus, and take corrective action." An example he gave of corrective action was an after-school tutorial program, funded through the curriculum division.

(5) Only one person expressed belief that teachers really utilize NRT scores. The director of junior highs felt that teachers got their fall 8th grade CAT test results back by Thanksgiving and that "that was plenty of time for second semester planning." He felt that "teachers can tell if they taught well by looking at individual students who they expect to do well, and seeing if those students do well."

The union representative, in contrast, said that "teachers find test scores of no value."

Attitudes towards NRT's.

a. By the community:

- "The public is complacent now that test scores are up."
- "The community is apathetic; there is no response, positive or negative."

b. By the Board:

- "When the scores were low, media coverage was high causing lots of Board interest; now that scores have turned around, there is no press and no interest."

c. By principals:

- "Principals are concerned about equity; e.g., my school has poorer kids and I won't look very good."
- "Principals are concerned about the use of test results in principal evaluations."
- "Principals have an informal competition among themselves based on test scores--e.g., tease each other in meetings about their school's scores."
- "The CAT is not challenging enough."
- "The bottom and the top of the CAT are soft; everyone is in the middle."

d. By teachers:

- "Teachers are concerned about the mis-use of tests for accountability."
- "They are concerned about too much emphasis on testing."

e. By central office staff:

- "We may overkill but there are vested interests in each kind of report."
- "Very positive attitudes, reflective of the results."

Criterion-Referenced Tests

- Math and Reading - Elementary Level

  (Given to all students in grades 2-6; been given for 6 years in reading and 7 years in math; items were developed by a task force of teachers; tests are given "to provide teacher diagnosis of student progress; teachers and principals use results through detailed progress reports required.")

- Math and Reading (optional) - Junior High Level

  (Given to all students; math just recently required [according to the director of junior highs, though the R&D unit person who filled out our testing chart did not mention this].)

- Optional CATC tests (Computer-Assisted Test Construction - criterion-referenced items at junior and senior high levels in the subject areas of English, General Math, U.S. History and Algebra.)

- State Proficiency Test

  (Given to all students in grades 9 & 11; subject matters include writing, reading and math; has been given in the district at the 9th grade level for two years and at the 11th grade level for one year; it was developed by State Task Forces; a passing score is required for graduation; other use included the expectation that "teachers and principals will use scores to identify areas of need, both individually and schoolwide.")

CRT Scoring/Reporting. The district requires a fall and spring administration of an "appropriate" level CRT for elementary math and reading, and for junior high math. A "CRT Package" is used to give the following kinds of information (via a computer-connected Scan-tron system):

1. District and school comparative data:

   Data is provided in the form of the school and district mean percent correct longitudinally and for the percent of students achieving above 60-70% for each sub-skill.

2. Frequency distribution by class.
3. Summary, item analysis.
4. Concept, item analysis.
5. Student scores list--individual student achievement--on each of subjects' major objectives.
6. Student response record:

   This is an item analysis by student.

As with NRT scores, the district attempts to get results back to principals for distribution to teachers within 3-6 weeks.

Proficiency testing results are provided by the following kinds of reports:

1. State/District/School Comparisons;
2. Student Profile;
3. Parent Notification;
4. Transfer Listings.

CATC tests are generated by computer, with answer keys supplied; teachers thus generate their own scoring and reporting systems.

Actual uses of the Criterion-Referenced tests. (1) CRT's are used as an integral part of a complete classroom instructional management system. Students are grouped for in-class instruction or placement into remedial classes based on their mastery or non-mastery of district-defined or state-specified objectives.

(2) Depending on the placement needs of students, CRT scores influence "schedule-building" of the junior and senior school level, i.e., the number and kinds of classes offered.

(3) Minimal competency test scores are used for communicating how the district's students are doing to the community, to the Board, and to parents. A good portion of a recent Board meeting was devoted to an analysis of these scores.

(4) CRT scores are used to pinpoint strengths and weaknesses in the district or school-level programs, particularly in content areas which are identified as needing more instructional time spent on them.

(5) The central office staff see the CRT's as a way of forcing teachers to pay attention to the district continuum. If teachers don't follow the continuum, their CRT's will "look bad."

(6) The R&D unit compares CRT scores to NRT scores to analyze course leveling or difficulty. For example, a recent examination of scores indicated that while 6th grade CRT math scores were very low, the NRT scores were "high" (in the 60th percentile). District staff are "looking at the 6th grade course to see if it is geared too high."

(7) Principals usually look at teacher utilization of the CRT instructional management system as a part of their teacher evaluation process. The actual scores are not supposed to enter into the evaluation, but teacher use of the system does. Principals, as well, are frequently evaluated based on their teachers' use of the system. (Principal evaluation based on teacher use of CRT's appeared to be emphasized more in schools with lower test scores.)

Attitudes towards CRT's.

a. By the community:

- "They're not concerned."

b. By the Board:

- [interested in minimal competency results]

?

c. By principals:

?

d. By teachers:

- "Some feel the district office is trying to take too much autonomy away."
- "Some think it's great to know what to teach, and how to make instructional decisions."
- "They make too much paperwork."
- "Some complain about having to teach to the test."

Perceived strengths of CRT's.

- "The state proficiency test will finally make some kids take testing seriously because it may influence their life chances."
- "Great organizational benefits."
- "Can improve average and new teachers."

Perceived weaknesses of CRT's.

- The focus on basic remedial skills to the exclusion of exposing or advancing students to higher level skills.
- Teachers may opt for easy topics (e.g., Ivanhoe, not Shakespeare) because their pupil performance will look better.
- The danger that teachers become "packet organizers" and lose their instructional interactions with students.
- The system focuses on individual skills, and in language arts doesn't allow time for enrichment reading or reading for fun.
- In language arts there is not really a clear hierarchy of skills, and the district interpretation of what should be taught may not be the best way.
- Too much information is presented; principals get computer printouts several inches thick.
- Can hamper "good" teachers; they will limit their activities to what the district wants and thus lose creativity.

The following chart summarizes the tests given and at what grade levels:

Chart 1

STUDENTS TESTED

	CRT reading	CRT math	Otis-Lennon School Ab.	SAT	CAT	Otis-Lennon Mental Ability	State Proficiency	CATE Computer Assisted Test construction
Grade 1	Optional	Optional						
2	✓	✓	✓					
3	✓	✓		✓				
4	✓	✓						
5	✓	✓	✓					
6	✓	✓		✓				
7	Optional	✓						Optional
8	Optional	✓			✓	✓		Optional
9	Optional	✓					✓	Optional
10		Remedial						Optional
11		Remedial			✓	✓	✓	Optional
12								Optional

CURRICULUM DEVELOPMENT

History

As was discussed previously, curriculum development in Crescent City has paralleled the organizational movement from decentralization and diversity to a more uniform and centralized program.

In the late 1960's the district had moved to a diversified curriculum in which each school had been encouraged by the superintendent to develop a program which would be relevant to its pupil clientele. It became increasingly evident that this was unworkable. A wide variety of materials was being developed and purchased. Parents were complaining that when their children were transferred from school to school they encountered different, and non-integrated instructional programs. With the appointment of superintendents who were more inclined toward centralization, the district began to shift in the early 1970's towards its present more centralized system.

When the present superintendent was appointed in 1978, he placed the curriculum directors and specialists under the associate superintendents for elementary and secondary instruction. This was an important change. By doing this, the district's curriculum expertise is directly applied to developing useful curriculum guides--and the activity is closely integrated with its ongoing instructional program. Teachers and administrators are vital participants in the curriculum development process. The resulting curriculum is more likely to be accepted and used by the teachers and administrators.

Curriculum Development Process

Perhaps this can best be explained by an illustration. For example, the decision is made to revise the business mathematics curriculum. The push to change this can come from anywhere--teachers, administrators, curriculum committees. The revision responsibility would be given to the curriculum department of the appropriate division, in this case the Superintendent for Secondary Instruction. A task force, consisting of business methematics teachers and chaired by a curriculum consultant would examine the present curriculum and write a new one. This might well take a year or more. Progress on this activity would be reported periodically to the respective curriculum director. The draft curriculum would be criticized by other business math teachers and then it is field tested for a year or more with all or a sample of district business math teachers. Appropriate revisions and adjustments are made and ultimately the final curriculum is accepted by the board and implemented. Once accepted by the board, business math teachers are expected to follow this curriculum. Their willingness and ability to do so is monitored through using the Elements of Quality and the instructional management plan.

The district curriculum guides are more thoroughly developed at the elementary than at the secondary level. A major district curriculum activity is developing comprehensive curriculum guides for all subject areas.

The curriculum specialists, all of whom are teachers on secondum, supervise the curriculum development committees and serve as curriculum consultants in the schools. As curriculum generalists, they visit schools when invited to offer assistance on implementing the curricula. In addition, they participate in developing and conducting the fall teacher orientation program.

The curriculum director meets weekly with all the division directors and at least twice a month, but usually more often, with the appropriate associate superintendent. The curriculum directors are advised by a principals' advisory committee and curriculum committees which consist of teachers, principals and representatives from the teacher association. In conversations with teachers, administrators and the teacher association it did not appear that these advisory committees played a major role in curriculum development. The main operational unit appeared to be the working committees of teachers and curriculum specialists who are presently writing district curriculum guides for the various subject areas and grade levels.

EVALUATION

The Organization of Evaluations: Who Evaluates, Where, What?

Who and where? The evaluation efforts seem to be divided between two groups of people: Those who initiate or oversee evaluations, and those who actually perform evaluations. Several people are responsible for initiating or overseeing evaluations. These people include:

(1) The Elementary, Junior and Senior High Directors: The Elementary Director stated that one of his many responsibilities was "the evaluation of elementary programs and the redirection of basic curricular substance and methodology." The Junior High Director looks for "skill areas I'm not satisfied with, and ways to redirect effort to remedy them." The High School Director tries "to evaluate everything we do, and anything we have changed."

(2) The Director of Federal Programs: Because of externally mandated evaluation requirements, must see that the required evaluations are performed and written.

(3) The Director of the Department of Research and Development: This Director gets requests from other administrators for evaluation information. He is responsible for ensuring that the information is produced.

(4) The Director of Special Education: Partly because of externally-mandated requirements and partly because of personal interest, she initiates requests for evaluation services.

Requests for evaluation services also come from other district sources. The sources for requests for evaluation information are not typically identified in the written information we have about the district. The R&D Unit's records, for example, are qorded passively (e.g., "a request was made..."). Apparently some come from the Board: A board memner mentioned that they can and do request evaluation or research studies from Bundren; a junior high principal remarked that the district research office does most of its work for the district administrators and the Board.

The second group of people associated with evaluation are those who actually perform evaluations. These people apparently are typically in the R&D department.

The following is an organizational chart of the district's Department of Research and Development and a brief description of each staff member's responsibilities.

Department of Research and Development
Staff Assignments

Dr.	Mr.	and Mrs.	Mr.	Dr.
Curricular Investigation*	Curricular Investigation*	Curricular Investigation*	Curricular Investigation*	Curricular Investigation*
District-wide Testing (K-12)	Special Needs School/Counseling	Johnson O'Malley (JOM)	Timesharing	Processing Center
CATC	Title I	Indian Education	Micro-computer Pilots	Divisional D.P. Requests
DAPP	Deinstitutionalized Handicapped	Executive Abstracts	SPSS	
DP/R&D Liaison	Comparability Report	Teacher Center		
Syllabus Testing	RIP	Rutland Center		
Reconstitution (Revision)	CRT Reading Investigation	OCR (TBE)		
	Structure of the Intellect (SOI) Institute Pilot	Fine Arts		
	PL-94-142	Title IV-B Library Sources		
		Title IV-C Minigrant		

*Department Priority

8/20/80

The Department's director has been in the district for over 20 years. His training and experience are mainly in school administration--he has been R&D Director for ten years. The number of department staff is small and has been declining somewhat during the past years. Beginning in 1969 with six certified and five classified portions, the Department now has five certified and four classified. There appears to be a contradiction here. While the district appears to be utilizing testing and evaluation more and more, the size of the R&D staff is declining. Also, given the responsibilities assigned, and the reported confidence in the work of the R&D Department, the Deparment does not exhibit the extensive staff training and background found in many other R&D units in districts of similar size. The two evaluators are teachers on leave; only the computer specialist and the testing director appear to have very much preparation in evaluation, testing and statistics. Yet this group is responsible for a wide range of evaluation, testing programs and special studies.

What is evaluated? The district seems to conduct three kinds of evaluations. One is the evaluation of specifically named and discretely identifiable programs, such as Title I, Indian Education and Title IV-C. A second, gathering information to assist in specific policy decisions. Examples of such evaluations are: The Length of School Day Survey, Credits Earned by Graduating Seniors, and ITV Summer Math Series Usage. These two kinds of evaluation efforts are formal, with written evaluation reports. A third kind of evaluation involves using testing information to inform decisions and make changes relative to general curricular emphases and methodologies. Ironically, this kind of evaluation is not written up formally, yet comprises the major evaluation effort and attention of the district.

According to the R&D Director, the following kinds of evaluations have had the most use in the district:

- a special study of the effects low pupil attendance has had on the district;
- various needs assessments carried out at school sites;
- a study of district SAT score patterns which identified that while the district was concentrating on raising minimal competencies, there had been a concomitant drop, as measured by SAT's, in the achievement of the district's more academically able students;
- reviews of the literature as requested;
- a study, now underway, to try to assess the productivity capacity of the California Assessment Program to predict student minimal competency achievement.

Evaluation Methods

The phrase "evaluation design" was not used in any of the interviews or in any of the district's written materials we received. Evaluations which gathered information for policy decisions tended to rely on questionnaires and/or existing records and documents. Evaluations performed to satisfy externally-mandated requirements used a simple pre post-test design, focusing on test score changes, generally over a school year. Evaluations of curricular emphases and methodologies based on test information informally used longitudinal pre post-test comparisons.

Attitudes Toward Evaluation

Attitudes seem to run from passive to positive, with no negative attitudes reported. The community apparently is fairly unknowledgeable and therefore passive about evaluation, except where there is a vested interest (such as advisory boards). In those cases, the attitude reportedly is positive. The Board expects evaluations and reports about specific programs. The Central Office attitude was uniformly reported as positive. Principal attitudes were reported as generally positive, while teacher attitudes were reported to vary, with Title I teachers very positive, and other teachers fairly neutral.

The future progress of this system is cloudy. Since the completion of our field work the district has experienced considerable turmoil. The state legislature which now provides a significant percentage of the district's budget, has implemented a fiscal austerity program. This, coupled with federal budget cutbacks, has placed a severe financial strain on the district.

One important effect of this has been increased tensions between the district's Teachers' Association and the school board and administration. Recently the school superintendent resigned, apparently because of his unwillingness to meet teacher salary demands and some school board members are facing possible recall.

While the district's instructional program does not seem to be the focus of the present situation, these externally produced problems will likely have a profound effect on the program's future. The main reason for this is that these successions and the uncertainties and turmoil that surround them, direct attention away from ongoing activities. Energies and resources are consumed in establishing alliances and setting up structures. In Crescent City this is happening just when the district is starting to implement its instructional renewal program in the secondary schools. Our respondents were aware of the difficulties of doing this under ideal conditions, and the need for time and resources in order to be successful. Likely, this present turmoil will slow, for the indefinite future, the move into secondary schools.

A new Superintendent has not yet been selected. The attitudes of the new superintendent toward this program and his/her commitment to further implementing it will likely be critical in determining future school district programs in instructional renewal.

BORDERTOWN

1981

BORDERTOWN

1981

BACKGROUND INFORMATION

The Setting

Bordertown is an older industrial city, with a population that has been declining for at least twenty years; most of the population loss has been of middle-class residents to the suburbs. The 1980 city population was 378,000, the metropolitan area population was 1,350,000 and the population within school district boundaries was 410,000. The city has a large number of Catholics, some Apalachians, and a gradually increasing percentage of blacks. As the city population has declined, so has the public school enrollment: from 87,500 in 1964-65 to 53,800 in 1979-80. The proportion of black students in the system has increased, from 45 percent in 1970-71 to 56 percent in 1979-80. Currently, about one-fourth of the school age children in the district attend private schools, predominantly Catholic schools.

Because of the declining enrollment, the school district has had serious financial difficulties, particularly in the last few years. Program and personnel cuts which the Superintendent called "drastic" were made in 1980. In June of 1980, however, Bordertown voters passed a tax levy which eliminated a projected $11 million deficit for the 1980-81 school year, and a possible state takeover of the school system. Paradoxically, the district is well-funded through federal programs. The district has received one million dollars per year in vocational services in recent years. Other federal and state funding comes through Head Start, ESEA (over $8 million), Disadvantaged Pupil Program Fund (over $3 million), Auxiliary Services to non-Public Schools (over $3 million), and Urban Education Pilot Project ($1 million).

The district operates 99 geographically-districted schools: 67 elementary, 2 middle, 13 junior high, 8 senior high, and 9 special (for physically or emotionally handicapped students). There are 13 alternative schools, which the district labels "undistricted", along with a college preparatory school and the 9 special schools. An open enrollment plan allows students to transfer if such a transfer will improve the school's racial balance.

Primary grades to grade 3 are organized, for the most part, into self-contained classrooms. Grades 4-6 are organized into middle schools with a "semi-departmentalized" program in which students are taught language arts and social studies by one teacher, math and science by another teacher, and P.E., music and art by specialist teachers. The secondary schools are organized on the quarter plan; hours of credit are awarded at the end of each quarter for each course.

Special education programs enroll 10 percent of the system's students. The district describes the special education program as "one of the largest in the state and most complete in the nation." The district provides several kinds of continuing education opportunities.

School District Organization

The organization of the school district and the relationship between the central office and the schools has been influenced by forces extended to the school system itself: historic, sociological and political.

Bordertown consists of many neighborhoods, some of which are identifiable by race, ethnicity and social class. These neighborhoods have a strong tradition of independence currently reflected in community organizations which deal with matters such as zoning, road construction and schools. School district officials accommodate as much as possible to local preferences, although tight budgets and the need to consolidate some neighborhood schools have recently come into direct conflict with this

The presence of a large Catholic population many of whom send their children to parochial schools, the existence of private schools for the Bordertown affluent and the escape-hatch of nearby suburban school districts, smaller and less problem-plagued, makes mobilizing city-wide support for public education a difficult task. Exacerbating the difficulties is the generally negative view of the public schools adopted by the media who attend to the problems rather than the successes of the system. The state legislature seems to share this orientation; it has the reputation of being pro-suburban and rural, anti-city in its allocation of the educational dollar.

The district has been responsive to these pressures--partly because many individuals in the central office are long-term residents of Bordertown. In contrast to other districts where planned rotation is common, principals are permitted to remain in their schools without transfer as long as there is no pressure from the community. Principals and teachers have had wide discretion in selection of text books. There are currently more than six reading series being used in Bordertown schools.

During the decade of the '70's federal monies were easily available to local school districts, and Bordertown's central office wrote many proposals for programs to assist special groups of students. They were encouraged in this by an active Research-Evaluation Director who provided assistance in the procurement process. Because different units within the central office assumed responsibility for administering the funds with particular subsets of schools or student populations, the availability of federal funds seems to have strengthened the district's tendencies towards multiple rather than single approaches to solving Bordertown's problems.

In summary, Bordertown has problems similar to those facing many large or medium sized urban districts:

- general population decline;
- shifting population among various sections of the city;
- increasing student transiency;
- increasing numbers of minority students, non-English speaking students or others who require special educational services;
- budget cuts;
- loss of public confidence in the public schools;
- blame by the public on school officials and teachers rather than on social factors;
- low morale on the part of some administrators and teachers
- perceptions of decreased student achievement as reflected by standardized test scores.

Bordertown has attempted to address these problems in a number of ways over the past ten years:

- work with the community to pass a levy increasing the local funds available to the district

- work with the media to refocus on the schools in a more positive direction;
- procurement of federal and state categorical funds;
- responsiveness to federal and state requirements designed by these policy makers and administrators to improve schooling;
- curriculum, testing and evaluation, staff development activities.

The remainder of this report describes district functions in these areas, describes the coordinating mechanisms among them and between the central office and the schools, and examines the effects of these activities at the school and classroom levels.

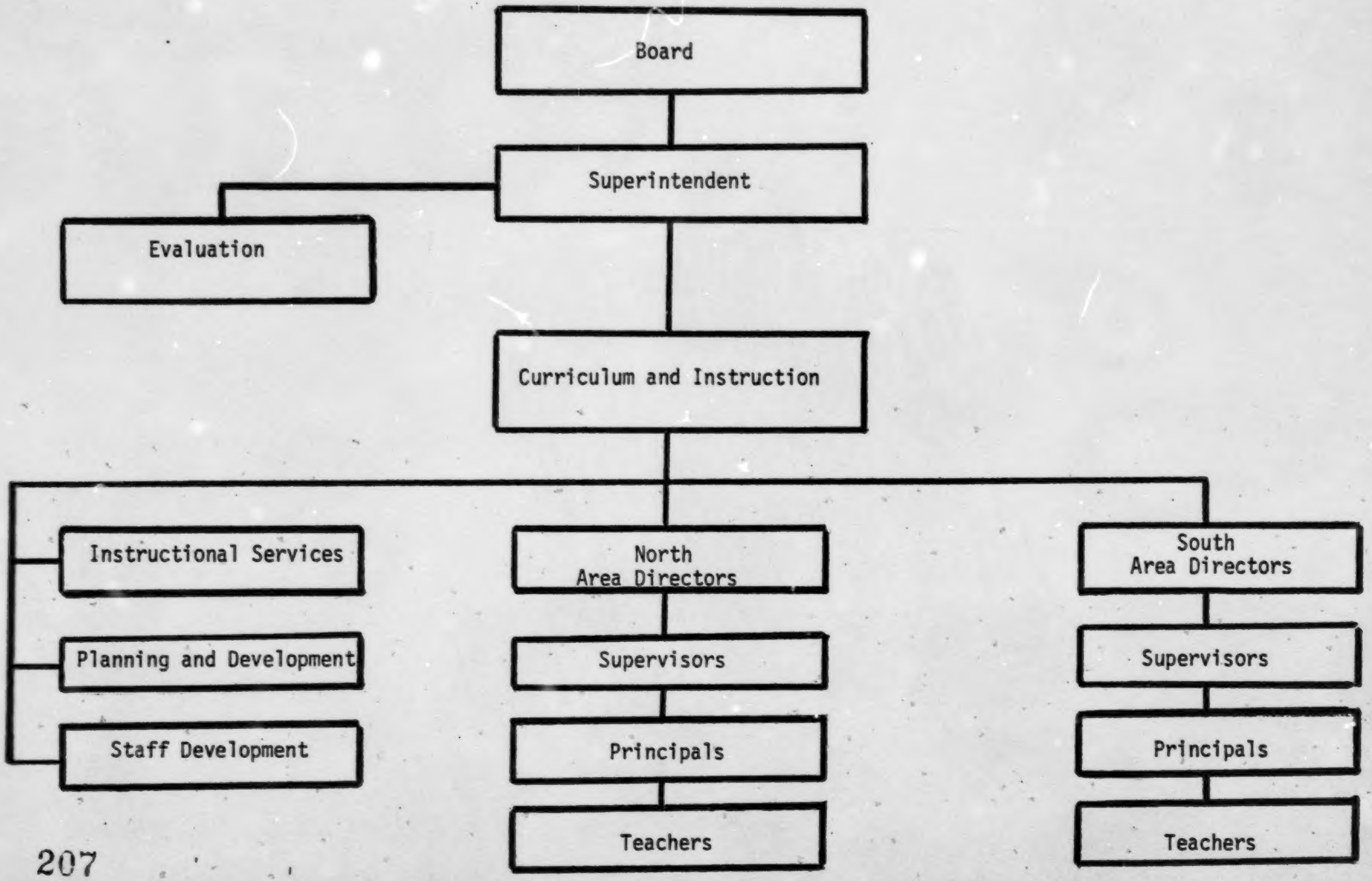
DESCRIPTION OF DISTRICT FUNCTIONS

The way the district divides responsibilities for major district functions is displayed on Bordertown's organizational chart (see Figure 1). The large Curriculum and Instruction Division includes an Instructional Services section responsible for doing curricular development; a Planning and Development section responsible for program development; and a Staff Development section. Program Evaluation, and its associated testing and data gathering, are located in a separate, independent division. This section of the paper describes the major activities of both the Curriculum and Instruction Division and of the Evaluation Division. It is based on interviews with individuals within these Divisions, interviews with principals and teachers as well as district-written documents provided to us by respondents.

re 1

BORDERTOWN PUBLIC SCHOOLS

Abbreviated Organization Chart


Board
Superintendent
Evaluation
Curriculum and Instruction
Instructional Services
Planning and Development
Staff Development
North Area Directors
Supervisors
Principals
Teachers
South Area Directors
Supervisors
Principals
Teachers

The Curriculum and Instruction Division: Instructional Services Branch

The Instructional Services Branch does curricular development. Until recently, the district's curricular content and its sequence were contained in pamphlets called Curriculum Bulletins. These pamphlets were written by Branch staff, and periodically updated. Content and sequence were both described in very general terms. The Bulletins were sent off to the schools. It was understood that principals and teachers were to read them. However, the extent to which they were utilized to control classroom instruction was not checked by the District staff. Principals and teachers did not refer to the Bulletins when queried about District curricular activities.

Beginning in 197 , a new "scope and sequence" was developed by the Instructional Services Branch staff. This document, called the Graded Course of Study, is a specific listing of minimal requirements in all courses mandated by the State, and provides a sequence for teaching skills in grades K-12. The impetus for the document came directly from the State; several years ago the State decided to enforce a statute which called for each school district to have a Graded Course of Study. The work on this document was described by one of the project directors as "a momentous undertaking," and as a "major thrust over the last three years." In his words, the Graded Course of Study is to become a "rather permanent" description of Bordertown's curricular content. Curriculum Bulletins are still in circulation, but instead of describing content, as they used to, they now provide suggestions for teachers, lists of reference materials, and other information which is updated periodically.

With the exception of a few vocational-education courses, the Graded Course of Study is now regarded by the district as complete. It has been disseminated to all schools. Central office staff are very conscious of the document. Most of our interview respondents in the district office mentioned it. However, none of our teacher or principal respondents in the schools talked about it spontaneously. When asked, one compensatory education director said that it was useful for new teachers, but 'others do their own thing." Since it is relatively new, it may be that principals and teachers do not yet react to it differently than to the Bulletin. On the other hand, it may be that the District office has not followed up with staff development or providing activities to indicate that compliance is expected. We were not told, spontaneously, of district efforts to ensure that teacher behavior occurs in accordance with the Course of Study.

The Curriculum and Instruction Division: Planning and Development Branch

The Planning branch was moved in the recent (1978-79) district reorganization from the Research and Development Division to the Curriculum and Instruction Division. The activities that branch staff perform, however, appear to be continuations of those performed before the reorganization. For example, historically and presently staff apply for new Title IV-C funds, and manage and evaluate IV-C projects which were previously funded. Five IV-C projects were in operation last year; four of them ended, while one continued and two new projects began this year.

Another ongoing activity of this branch has been the development and refinement of a criterion-referenced diagnostic testing program called the Bordertown Instructional Management System (BIMS). Over eight years ago BIMS was started with Title IV-C funds. The development process began with a review of the five to six different tests which were in use in the schools at that time, and the identification and positioning of specific skills where the texts devoted the most attention to them. Teacher committees then wrote items to test the identified skills. They began with math, then moved on to reading. The process anticipated for using BIMS included student testing at a frequency determined either by a teacher or the school, about once a month. A computer would be used to score and then report out the scores of individual students. The program director told us that teachers were supposed to use BIMS scores to "track progress through skill hierarchies."

BIMS during 19 was offered to schools on a phase-in basis. The district hoped that school principals would volunteer to have their teachers use BIMS at the rate of twenty new sites a year. But as volunteers slowed down, "some schools were pushed in to meet the twenty." Though there has been some talk, even at the School Board level, of eventually requiring all district schools to use BIMS, apparently this has not yet become outright policy. Now, there appears to be less rather than more central office control over BIMS adoption and utilization. The program director told us, "We exercised far more control of it in the development than we're doing now; you can't ram things down people's throats." Principals, teachers and program directors corroborated this statement; many saw less use of BIMS now than in previous years.

However, there is much activity currently going on in the BIMS project. Some involves converting the system to a new computer. Other efforts are directed toward writing new items that "flow from" the new Graded Course of Study, so that BIMS will be better correlated with content that is described by the district as "minimal." A third effort involves correlating BIMS items with newly developed minimal competency items and skills, and with the newly selected norm-referenced achievement test. Additions to BIMS include coding of texts to BIMS, and the development of a "career component" where the same BIMS skills are presented in a "work context." The criterion-referenced testing system that is BIMS has been in the works for eight years and only now are the connections between texts, tests and curriculum being made. The computerized feedback procedures to provide teachers and principals with results in a form useful to them are not quite in place. Neither is the staff training nor supervision which seem essential to make the system a useful diagnostic instructional tool.

Another major activity performed by the Planning and Development staff is development of a minimal competency testing system. Work on this system was initiated because of a motion by a Board member, not as in some other districts because of a state requirement. The Board called for standards to be identified, and then for students to be required to demonstrate competencies prior to grade promotion. Originally the Board wanted this to occur at every grade level. Branch staff, however, convinced the Board to focus on only grades 3, 6, and 9. Additionally, the Board has accepted an administrative recommendation that students who fail the tests will not be retained, but instead will be placed in remedial programs in their next higher grade level.

The district then applied for and received Title IV-C money to develop the tests. The money had a string attached to it, however; the state required the addition of grade 11 to the system. Branch staff conducted some preliminary work on the system at grade 11 and found that there was great teacher resistance to upper level testing. Additionally, staff "felt there wasn't adequate remedial effort which could be made so late in a student's career." In the light of these findings the district eliminated competency testing at grade 11 without penalty from the state.

Tests have been developed at levels 3, 6, and 9. Test kits are sent to each school. Kits contain: a practice form of the test; a teacher's manual and ditto masters for some related instructional activities; and a final test form.

Up until 1982, utilization of the system will be strictly voluntary. Twenty-four schools used it in 1981. Beginning in 1982, third grade testing will be required district-wide. Decisions have not yet been made about when to require testing at the other levels. The project director expects that about 50 percent of the schools will voluntarily use the system next year, claiming that there is a "desire for it because it is not just testing but also is a program."

Another major effort of the Planning and Development Branch is the School Improvement Program (SIP). SIP is directed by the head of Planning and Development; however, in addition, staff from other branches, including the Evaluation Branch, are on the committee which plans and oversees the implementation of the program. (This is the only program we observed in which staff across branches are working together.)

There are three components to SIP. One is "overall awareness" in which all district schools are asked to focus on instructional improvement. Planning and Development staff created three "how-to" telecasts on the theme of improvement which were broadcast over the educational television channel that school staff are encouraged to watch as a group. The telecasts focused on classroom organization and management, such as time on task; on discipline; and on research-based elements of effective schools. The head of P&D commented that the last topic may "cause a big row because it contains checklists to fill out to check teachers."

A second component of SIP are studies by Evaluation Branch staff. They are looking for factors which distinguish schools where student scores are declining from those where scores are stable or going up. Preliminary analysis shows that one significant variable distinguishing declining schools from others is high teacher and principal turnover.

A third SIP component is direct intervention. Project staff looked at schools with more than one NCE loss from the district average. Twenty-five schools in this category were invited to participate in developing a process to develop a "model for intervention." Ten schools volunteered for participation, and six were selected. Principals and two teachers from each of the six met with project staff for two full days. They "went through the whole process of looking at data and what it means, focusing on standardized achievement test scores." Each team built plans, which they took back to their total school staffs. Now they are working on "action plans." The project director expects changes to begin occuring in these schools over the next 1½ years, but anticipates that test scores may not reflect the changes by then because of the recent change in the norm-referenced test given to district students (from the Metropolitan to the California Achievement Test). He also anticipates that as the schools wrestle with these issues, they will begin to realize the district has some things to help them, expecially the "new BIMS, and the minimal competency tests."

The Curriculum and Instruction Division: Staff Development Branch

"We are a clearinghouse for in-service. The in-service in this district is a school-based effort. We are here to assist in that effort." This quote reflects the point of view of those in the staff development branch.

The primary function of the staff development branch is facilitating, arranging and doing the paperwork connected with the presentation of staff development courses for teachers and principals. The remainder of the staff time is spent in "coordinating"--that is, attending meetings, writing reviews and plans, etc. The Branch is service oriented; it has no program of its own.

Staff development activities have borne part of the brunt of uncertain district funding. Up until this year, staff development duties were coordinated by a single person. This year, there are 2½ people. Next year, there is no telling what the staff allocation will be. This year, an important activitiy for the Branch was a needs assessment formulated with the help of the Evaluation Division to determine what assistance teachers and principals most desired. ("Techniques for motivating students" came out first on both teacher and principal surveys while "options for dealing with stress" came out a close second.) There is no assurance that the district will be able to act next year on the findings.

Up until this year, most staff development activities have been squeezed into staff meetings, occasional workshops and university courses pursued independently by teachers and principals. Some special courses in intercultural relationships have been offered to the district by an outside university. Foundation funds have made possible management training courses for mid-management administrators and supervisors.

Staff development activities in Bordertown are not planned specifically to reinforce the District's thrusts in curriculum, testing or evaluation. Rather the activities occur in response to teacher and principal requests; or in response to requirements written into categorical programs.

The Curriculum and Instruction Division: Area Directors, Supervisors, Principals, and Teachers

Schools and their associated administrators are divided into two geographic groups; each with a "line" structure consisting of two area directors overseeing principals who, in turn, oversee teachers. In addition, supervisors of subject matter work with teachers and principals on such matters as text adoptions and curriculum, and in-service training.

Area directors both supervise and provide service to principals. Problems or questions that a principal may have should be taken directly to the area director; the area director handles what he can alone, then takes those he can't handle to the Associate Superintendent of Curriculum and Instruction and his staff. Possible solutions are discussed, and the area director communicates the information back to the principals. One area director reported that he and each principal "work as a team to deal with the problems we've got." This same director estimated that about 60% of his time was spent working directly with principals, while the other 40% involved everyday administrative duties to keep things running.

The interaction between principals and area directors is not always initiated by the principals. At times, the area director initiates contact. For example, a director may contact a principal because it is "my job to go in and find out why test scores are low, what happened." Directors use achievement test scores "as a way of discovering problems with the system." One director told us that test scores serve "as a bell to make me look further into certain analyses of school level problems identified by this approach: substitute teachers with no lesson plans in classes for a major part of the year; special education students mistakenly included in the tested population. Directors also use achievement test scores in goal setting meetings with principals.

Another responsibility of the area director is to appraise principals. Until recently the district used a form for principal appraisal every several years which simply asked for the principal's goals and documented the number of director visits to principals. A new state law requires that every administrator must be appraised every year. Bordertown, therefore, is developing a short form for principals who "we feel are coming along pretty well." For others, who "need extra help, especially the newer principals," the former longer form will be used. Principals are not evaluated by their students' test scores, nor by their teachers' behavior or activities.

Teachers are appraised by their principals. Area directors do not typically get involved in teacher appraisal; they do, however, "help principals if the time comes to get rid of a teacher." Teachers are carefully appraised "in their first couple of years before they get tenure." The approach utilized is to work with teachers to try to help them if they are not doing well. Teachers are appraised on a number of things. The Associate Superintendent of Curriculum and Instruction told us that, "We have assured teachers that the system [i.e., BIMS test scores] will not be used for teacher accountability purposes; however, some of them don't believe us. And we have also said that the appraisal of a teacher involved many data sources. The principal certainly has the right to look at the scores of students on the BIMS when he makes an assessment of teachers."

Tenured teachers are rarely fired. Last year, the district tried to fire two tenured teachers, and failed; they lost both cases in arbitration.

The Curriculum and Instruction Division: Summary

Major activities of the Curriculum and Instruction Division include the following: 1.) The Graded Course of Study and related Curriculum Bulletins; 2.) alternative schools; 3.) Title IV-C projects; 4.) Bordertown Instructional Management System (BIMS); 5.) Minimal competency testing; 6.) School Improvement program; 7.) Staff development; 8.) problem-solving at the school level; and 9) principal appraisal.

The Curriculum and Instruction Division turns out a lot of work and is involved in many different projects. Some staff seem to know about projects other than those they themselves are engaged with. Staff in other branches seem to be working independently. There did not seem to be a strong sense, among our respondents, that their jobs were part of a coherent and comprehensive instructional renewal program.

The Evaluation Branch

In the recent district reorganization, approximately two years ago, evaluation was moved from branch status within a Research and Development Division to independent branch status, with the Evaluation Director reporting directly to the superintendent. The Evaluation Branch is currently divided into four sections: testing, program evaluation, communications/reference, and school information. Discussions of activities associated with each section will follow. Activities relating to a developing decentralized budgeting program will also be described.

Evaluation Branch: Testing Section

Founded in 1938, the district's testing branch began by giving aptitude tests in grades 3,6, and 8. This annual practice has continued through to the present, except in 1977 when a teacher's strike prevented all testing. Achievement testing in grades 1-8 was added later. Until 1972 testing was originally administratively placed in the Curriculum and Instruction Division. In that year, when the current district superintendent became head of the newly formed Research and Development Branch, testing was moved from C&I to become a division within his R&D unit. In the recent district reorganization, testing was demoted from separate branch status in the R&D Division, to section status within the Evaluation Branch.

The Testing Section is responsible for administering the system-wide norm-referenced tests given in the district. These tests include: The California Achievement Test (grades 1-8); the Otis-Lennon Ability Test (grades 3-6); a selection test for 6th graders who want to enter the college preparatory school; and the GED test. Testing staff also administer various ESEA instruments, some of which are attitude surveys and some of which are aptitude tests. Staff additionally does some testing for the Advance Placement Program. Responsibilities included in all these testing activities are scheduling, coordinating, administering and overseeing the scoring and reporting through data processing.

Major current activities in which testing section staff are involved include those which are associated with changing the norm-referenced achievement test given to all students in grades 1-8. Until 1981, the district used the Metropolitan Achievement Test (MAT). In 1981, however, the district decided to switch to the California Achievement Test (CAT). Our respondents generally agreed to the reasons for the switch in tests: the content validity of the CAT is higher for the district's new curricular scope and sequence as delineated in the Graded Course of Study. Some others said, "It was time for a change. We were getting too used to the MAT."

Test staff are pushing for the change in the way scores are reported: the district has begun to use normal-curve equivalents (NCE's). While previously, scores for all Title I schools were reported in NCE's, this form of reporting is now being expanded to include all schools. Area directors, coordinators and principals are being oriented to these score interpretations by testing staff.

Another current emphasis is solving the problems rising from computer system changes. The conversion to a new system is considered especially significant because, for the first time, the district will be able to efficiently report scores for individual students.

Evaluation Branch: Program Evaluation Section

A large part of this section's effort is supplemented by funds from Title I schools. Section staff conduct Title I evaluations. They organize meetings, provide school staff with progress reports, and write up reports for submission to appropriate agencies. Title I evaluations are routinely conducted according to federal guidelines: program students are coupled with non-program students on pre- and post-achievement test scores and attitude surveys. Data from these evaluations are also used informally. A unique system has developed whereby Local School Evaluators are assigned to schools. This individual prepares data for local schools' use. This may mean preparing charts or graphs for specific groups. For example, parent opinion survey data is of interest to parent advisory groups. Local school evaluators help lead teacher meetings to analyze scores to determine what went well and what did not, at the school level. Other group meetings analyze the scores to determine what did or did not go well at the program level.

For example, several groups met to look at "a major needs assessment of the Title I program at the intermediate school level to find out why we were not getting gains there that we were at other levels. We brainstormed every possibility to what could be happening out there, from the composition of the Title I teaching staff to inconsistencies of effort as funding runs out." The following changes were made: new materials were purchased, teacher in-service programs were changed, and Local School Evaluators were assigned to conduct problem-solving group sessions with staff at the schools involved.

The Evaluation Section has four staff members (in addition to the section head), each of whom works at least part-time as a Local School Evaluator. The person in this position, originated over seven years ago by the current Branch director, "serves as a liaison between the branch and the schools." Most of the Local School Evaluator (LSE) effort is funded by and associated with Title I schools, although a small amount of general funds are allocated for LSE work in non-Title I schools. Originally, LSE's were conceptualized as "a kind of consumer information service that principals could call upon at will." Not many schools took advantage of this service, however. One of the problems might have been that principals are swamped with things to do, and "they just don't have time to sit back and reflect about things." Another problem was that principals did not have any background in how a local school evaluation might help him/her to solve school-level problems. LSE's do attempt to bring information to each principal, "in interpretable form," to at least get him/her thinking. At times, when the principal requests it, data is also presented to teachers. Long-range planning for LSE's is difficult because of their funding dependency on Title I; no one can predict from one year to the next how many LSE's there will be in any future year.

The Local School Evaluator is an interesting concept which makes evaluation services at least potentially available to schools. It is clear that many principals did not understand the need for such services. Other principals do, in fact, solicit data-based information and have incorporated their Local School Evaluator into both formal and informal planning and troubleshooting. Others barely know the name of the person assigned to their school.

Conversely, Local School Evaluators because they cannot serve all their schools adequately, have selected certain principals or schools for which they have special affinity. In these schools, their services--often adjunct to their coordination skills--in helping to resolve difficulties and suggest alternative services are much appreciated.

Although the conceptualization of the Local School Evaluator is an invention to the credit of Bordertown, the institutionalization of the system is not yet evident; rather, when it works, it seems to work because the skills and interests of the local school evaluator, as a person and as a professional, seem to mesh with the professional skills and personal characteristics of the principal.

Evaluation Division: Communications/Reference Section

This small section is primarily responsible for publishing materials targeted to different audiences. The most common publications are called branch reports. Another is an educational periodical designed to report research and evaluation findings "in some depth." This periodical is carefully written in language that can easily be understood "by most persons."

Evaluation Division: School Information Section

The School Information System (SIS) was begun with Title III money in 1970. The project director described the major goal of this effort as providing decision-makers "with relevant, timely, reliable, and valid information, presented in an easy to read fashion." Data is delineated, gathered, analyzed and reported using the school as the basic unit of data aggregation. Individual student or class information is not provided. More than 800 variables per school have been collected and reported every year. The categories of variables included: pupil, such as attendance, achievement, and attitude; staff, such as attendance, experience, and attitude; school plant, such as rooms in use and capacity; costs, such as per pupil and per school; demographic characteristics, such as parent attitudes, incomes and education.

Much of the information used to compile the SIS is collected by other departments. The only new data, originating yearly, from the section itself are surveys of student, teacher, parent and administrator attitudes. Among the major reports which are generated yearly are: 1) an exceptional characteristics report, in which variables which correlated with student achievement variables were identified; 2) a variable print-out; 3) the specific results of the attitude surveys; and 4) a trend report, in which values for selected variables were graphed over the five previous school years.

The information from the SIS is often used by the local school evaluators when they go out to work with schools in their "planning for the next school year" capacity. SIS data are used, also, to display trends to the public in a variety of district-written publications, as well as to identify district-wide problems needing attention. For example, recent declines in student and teacher feelings of safety resulted in new security measures being taken.

Evaluation Branch: Local School Budgeting Liaison

Bordertown began experimenting with a new approach to school site budgeting last year, in its program called Local School Budgeting (LSB). The approach has two main characteristics that distinguish it from traditional budgeting procedures: community participation and decentralization. A publication about the program described the approach as "predicated on the conviction that those who pay for, participate in and benefit from the schools should have a strong voice in how school revenues are spent. School goals and the expenditures to reach these goals are set at the local school level. "

The program was stimulated by active citizen participation in the passing of the tax levy. It began last year when a Steering Committee was chosen by the community. The Steering Committee (which included the head of the Evaluation Branch as the superintendent's representative) selected seven schools to participate in the program's first year. Another three were added to the program through a grant from a local organization. Teams were formed in each of the ten schools; each team included parents, community members, and staff.

Team members participated in several training sessions. Training centered around finance and budgeting, but training in educational planning and group process skills was made available to members who desired it. Each school was then allocated money derived from a traditional method of resource allocation, based on such factors as enrollment and special programs. The budget teams then made recommendations on how the money will be spent, within constraints imposed by the Board, teacher contract, state law, etc.

Three Evaluation Branch staff members were assigned to serve as liaisons to the budget teams from the central office. Their responsibilities include communicating information to and from the teams and securing training or consultation help from other central office personnel.

The program is rather small now, but is expected to grow over time. The Board recently passed a motion calling for the establishment of local school advisory committees in all schools. The current budgeting teams, then, are seen as the prototypes of these advisory committees. The Evaluation Branch director told us that they "hope to have this program institutionalized in all the schools within five years."

Some of the school teams have been "quite successful," while others have not. Reportedly, some opposition to the program has come from principals who see the teams invading their decision-making territory and consuming a lot of time without much payoff. Another major concern lies in the as-yet-untested question relating to whether the schools will be able to implement the changes that the teams want at the school site, when those changes may conflict with usual district practice or policy.

Evaluation Branch: Summary

Major activities of the Evaluation Branch staff include: 1) norm-referenced achievement and aptitude testing; 2) Title I and other program evaluations; 3) Local School Evaluator test interpretation and in-service training for individual schools; 4) publishing and disseminating reports and periodicals; 5) compiling and disseminating the School Information System; and 6) working with the Local School Budgeting program.

WHAT BORDERTOWN IS DOING TO LINK TESTING AND EVALUATION WITH INSTRUCTION

Two observations about what Bordertown is doing to link testing and evaluation information with instruction: First, there is district interest in focusing attention on lower-scoring schools and students. Second, is an emphasis, not on a single comprehensive, coordinated approach to improving instruction but a preference for separate multiple lines of effort. Each of these will be discussed below.

District Emphasis on Lower-scoring Schools and Students

There are several ways in which the district focuses attention on lower-scoring students and schools. A formal way is through the School Improvement Program. In SIP (described in the previous section as an activity within the Curriculum and Instruction Division) central office staff make direct interventions with schools with declining test scores. A school is considered a "declining" school if student scores on the district norm-referenced achievement test are more than one normal-curve equivalency (NCE) below the district average. A number of schools meeting this definition are analyzing their achievement score data and developing school-based plans for action.

A related, though less formal, approach is used by the Area Directors (those line administrators who supervise principals). When they see particularly low test scores coming from their assigned schools, they conduct individual investigations with the school staff to attempt to identify reasons for the low scores.

Information on student and staff attendance and turnover, as well as student motivation and discipline, are gathered and analyzed. Reasonable actions are taken with the goal of improving student test scores.

A third way in which the district focuses attention on lower-scoring schools and students is through the district student placement process. Bordertown's policy governing student placement into classrooms is to avoid so-called "dummy-tracks." If individual students are having problems in a class or a subject matter, they are scheduled into special remediation for a particular time period. Thus, a student with reading problems may be assigned to a reading laboratory for an 8-10 week session. A student with broad needs may be assigned to a particular compensatory education program; here, as with the shorter-term remedial programs, the goal is to bring the student "up to speed" and return him/her to the classroom.

These ways of focusing on lower-scoring students and schools have two elements in common: 1) students and schools are identified as needing attention, primarily based on achievement test scores--criterion-referenced scores are not used; and 2) intervention tends to be taken at the school level, not at the class, program or district level. Low test scores serve as a red flag, a warning that something out-of-the-ordinary is happening. Scores are not typically used to identify specific subject matter instructional needs of individual students, nor to implement district level instructional changes.

Promotion of Separate Multiple Lines of Effort to Improve Instruction

Some of Bordertown's many central office activities related to instruction were described in the previous section. Most of the central office's specific programs originate in two branches: Planning and Development, and Evaluation. They include: Bordertown Instructional Management System (the district's first criterion-referenced system); Minimum Competency Testing; School Information System; Local School Evaluators; Norm-referenced achievement testing and the School Improvement Program. Other central office activities include the Graded Course of Study and Curriculum Bulletins.

These activities share one characteristic. While most attempt to link testing and evaluation information with instruction in imaginative and productive ways, they do not appear to be part of an overall comprehensive coordinated district plan for improving instruction. Instead, each is a separate independent program, presented to individual schools in a smorgasbord of offerings. Not only are the programs separate and independent, but each tends to maintain and guard its independence. Each begins with a constituency and works to maintain and expand it.

There are several potential explanations for this multiple thrust approach. One is that the central office maintains a responsive, rather than an pro-active, orientation towards the schools. This may be because this is how Bordertown has always been, or it may be a result of the individual predisposition of those in key roles.

There seems to be some tension within the district between tendencies towards decentralization and those towards centralization. The declining school population making school consolidations a necessity combined with the district's financial difficulties, lead the district away from community-based involvement, and towards centralization. So does the development of the Graded Course of Study which has created a new trend toward aligning some of the other previously independent instructional components with it and with one another (i.e., BIMS, minimal competency, and the norm-referenced tests). This trend toward alignment supports centralization, as it represents the beginnings of an integrated district policy. However, the historically grounded independence and competition between central office groups was maintained through the recent district reorganization, when the Superintendent kept the Evaluation Branch a separate branch reporting directly to him. This separation tends to support decentralization; the Local School Budgeting movement supports decentralization. And, cutbacks in the administrative staff because of the financial problems also support decentralization, as there are fewer monitors to promote compliance with district policies.

The district's multi-thrust approach allows the district to keep operating even though the tension between decentralization and centralization is not resolved.

An additional factor which may explain Bordertown's multi-thrust approach is that the district has several kinds of political pressures to deal with. One is the uncertainty and unevenness of funding from one year to the next. Many of the district's programs are federally funded; some elements of these programs operate only from year to year. Staff do not know which of these program parts will continue in the following school year until late in the current year. Also, the district's general funding is very uncertain. A tax levy passed last year should have provided funding for operations for several years; however, the state has cut back in their funding support for the district, so the district is still in great financial difficulty. This makes planning for an integrated, coordinated district program difficult.

Another political pressure the district is feeling is the desegregation litigation which has remained unresolved for many years. It is possible that district staff may not want to promote a coordinated district program which could possibly be used against them in this ongoing legal case.

There are factors present in the district, therefore, which make the district multi-thrust approach understandable. Perhaps this approach is the only possible one, given the current district situation. However, stepping back from the realities of the situation, the implications of this approach seem not to be positive for long-term incremental instructional improvement. Without more coordinated management planning, the sum of the parts may be less than the parts themselves. Without a district vision which is clear, understood by all schools in the district, and whose implementation is encouraged and checked, classroom-level instructional improvement will continue to be primarily dependent on the presence of extraordinary teachers, principals, and administrators.

COORDINATING MECHANISMS

Bordertown does not appear to have a unified approach to instructional change. Concomitantly, the district seems to have few formal linkages to maximize staff committment to and staff communication about data-based instructional change.

However, there are some formal committees composed of members across central office units, and some which include field and community representatives. For example, there is a community communication council, made up primarily of business people and some parents. This council has recently looked at the competency program process and communicated about it to the community. Also dealing with the competency program is the Competency-Based Education committee just being formed, composed of staff and consultants associated with the competency-based movement. The head of the Planning and Development Branch chairs a planning and dissemination committee, whose members include the Evaluation Branch head, and Curriculum and Instruction and Title I staff. The Committee's purpose is to coordinate activities across the various organizations; members meet about four times a year. Two of the central office programs (the School Improvement Program and the Local School Budgeting effort) have committees composed of staff across central office divisions, and in the case of the LSB, also composed of school staff and community representatives.

Another kind of coordinating mechanism is currently being installed in the district. The new computer system will allow scores and other information about students to be listed by student identification, for the first time.

An informal coordinating mechanism lies in the staff interactions with the School Board. The Board, which often originates mandates for activities, is informally "educated" by committees and individuals which meet with Board members periodically. For example, the Board originally wanted every grade level to be involved in the minimal competency testing, and also wanted promotion to be based on the yearly demonstration of competencies. The Planning and Development Branch, however, promoted the testing of only grades 3, 6 and 9, and also promoted the use of test scores for remediation but not promotion. The Branch head told us that the branch "influenced" the Board through "our education of them."

EFFECTS

As might be inferred from the previous discussion the extent to which the district's approach is influencing instruction as it occurs in classrooms varies from school to school. Many of the central office's instructional programs are intentionally offered to the schools for their voluntary utilization. But even those programs in which all schools must participate, such as the norm-referenced achievement testing program, have an unpredictable effect on instruction, as the results are used differentially throughout the district. Program by program effects will be discussed in this section, as well as the effects of the policies on the district as a whole.

Effects: The Graded Course of Study

Although district office staff frequently mentioned the Graded Course of Study, none of our teacher or principal respondents talked about it. One compensatory education director said that it was useful for new teachers, but "others do their own thing." Perhaps one reason the course of study does not appear to be considered an important instructional component in the schools is that not much attention appears to have been placed on making teachers and principals aware of it, training people in its use, or ensuring compliance. Thus, its effects at the school and classroom levels at the moment appear to be very small.

The effects of the course of study at the central office level however, may prove to be great. Because the district has been forced to describe its curricular scope and sequence publically, the course of study is being viewed as the focal point for an alignment of some of the district's instructional offerings.

Effects: BIMS

Use of the BIMS criterion-referenced diagnostic testing system is at the discretion of individual principals and teachers. Estimates of school-level usage reported to us ranged from "all but five schools" to "out of my assigned 25 schools, only two or three use it consistently." Respondents also talked of differential rates of use of the system by teachers within schools. Teachers in laboratory settings (where much of the district's remedial effort is concentrated) reported higher and more consistent usage than did regular classroom teachers. Reported use extended from elementary school grades to grades 7 and 8, with use at the higher secondary level reported only in association with a compensatory education program.

Many reasons were given us for the apparant low-levels of interest in BIMS. Principals and teachers talked of using other systems which conflicted with BIMS, such as the Fountain Valley Diagnostic System in Title I schools, and the introduction of a new management-oriented reading program which cannot be operated simultaneously with BIMS. Principals also spoke negatively about the cost of BIMS (costs associated with computer link-ups), and about the past lack of correlation between their curriculum and BIMS items.

Central office staff reported other concerns which inhibited use of BIMS. These included: teacher fear of using a computer; long turn-around time; system complexity; and questions about whether using the system conflicts with "humanistic education." Evaluation office staff had three kinds of comments about the problems of BIMS. Initial student placement is a first concern. Reportedly, students who "have the same g.e.s will have 2.5 to 3 years range on the BIMS." If a student is placed two or three years too low, they will make some misleading apparent gains as they take the system's tests. A second concern is over the number of times BIMS tests are given in a school. We were told that some schools test only once a year, while others test 8 or 9 times. "If a student takes only one test he is only going to go up one level, even if he's a genius." A third concern is about expectations as to what constitutes a "good performance." No one has looked at BIMS in light of setting standards.

Another factor which probably influences non-utilization is the lack of a comprehensive staff development effort in this area. The program developer told us that they used to train principals about BIMS, and expected the principals to train their staff. Now they are trying to work together with principals and supervisors to train the staff.

In spite of all the work on the system, it does not appear that the use of BIMS is enthusiastically encouraged. Evaluation Branch staff are contributing none of their expertise in addressing some of the system's weaknesses, but instead openly express their negative opinions about the system.

Since the use of BIMS is irregular, its effects at the school level are likely to be mixed. Effects at the central office level seem minimal, with only project staff working on it, and no other staff using the system or its results in any of their work.

Effects: Minimal Competency Testing

Because minimal-competency testing has not yet been required of all schools across the district, no effects were reported to us. Predictions about possible future effects were made by central office staff. These predictions ranged from doubts about the testing making any difference unless it was used for determining student promotion to assertions that the testing will have major impact on instructional content, even through the secondary level.

Effects: School Improvement Program

Our school-level respondents did not discuss the SIP with us. Central office staff, however, expressed much interest in SIP. This may be because SIP is the one program in which central office staff from several branches are all participating together. Attitudes are generally supportive of the program, though its effects are currently unknown or unreported.

Effects: Norm-referenced Testing Program

This is one district program which is required of all schools, grades 1-8. Teacher and principal attitudes toward the test itself varied, with some neutral comments and some negative. Neutral responses indicated an understanding of why such tests are required, and an acceptance of them as a necessary evil. The negative comments focused on their lack of relevance to teachers (e.g., results are received late in the school year, the information received is not of much use to their instructional planning), and on their cost in terms of time taken away from other instructional activities.

Utilization of information from the testing varies from school to school, depending on the beliefs of principals and teachers, on what particular district sponsored programs the school is involved in, and on whether the scores are lower than expected. Some principals seriously use the scores in their goal-setting and planning activities and pull their teachers into using them. Title I schools in particular use the scores for program analysis and reporting purposes. And NRT scores are used by area directors and SIP program and school staffs as signals to look for school-level variables which may be depressing student achievement and test scores.

Effects: Local School Evaluators

Since LSE's are primarily associated with Title I schools, most "regular" teachers are either not aware of them, and so have neither nor negative attitudes toward them. Title I teachers seem to be affected differentially, depending on whether their principals understand and support LSE activities. The major use of LSE's currently is in test score interpretation. One of the LSE's said that "if it weren't for LSE's, the teachers wouldn't really get into test information; the principals might not unless there were problems."

Effects: School Information System

Every school receives SIS information; however, as with the norm-referenced testing activities, utilization of SIS information varies across schools. Many of the school site respondents think of SIS primarily or solely in terms of the attitude surveys which are conducted annually for the SIS compilation of information. Actual use of SIS information for planning or assessment at the school level by principals and teachers was described to us infrequently.

Effects: Local School Budgeting

The Local School Budgeting effort is a relatively small and new one. Specific effects of the program on the participating schools were difficult to identify. Central office staff reported that success varied greatly across schools. At this point in time, we did not detect that the program was having much impact on the central office either, other than on the staff directly responsible for period visits. However, because the program obviously represents a decentralized approach, the program's future success or failure may influence the central office tendency to go one way or the other on the centralization/decentralization issue.